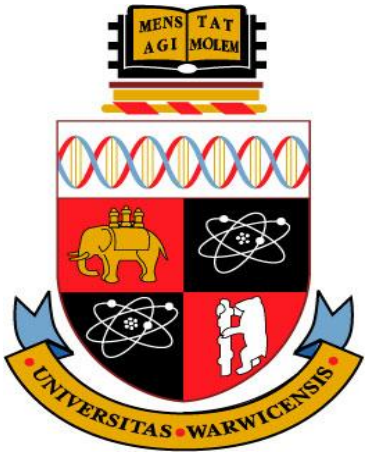
Investigations of Siderophore and Tetronic acid Biosynthesis in *Streptomyces scabies* 87.22

Joanna Bicz

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Chemistry.



University of Warwick

Department of Chemistry

March 2013

Contents

List of Figures

List of Tables

Acknowledgements

Firstly I would like to thank Professor Greg Challis, who gave me the opportunity to work on this interesting project and for all his help and advice throughout the whole period of my PhD. I would also like to thank all the members of the Challis' group for their support and creation of great working atmosphere. I would also like to thank Dr. Paulina Sydor for her guidance throughout duration of my studies, especially during the first year. To Dr. Christoph Corre, Dr. Nadia Kadi and Dr. Orestis Lazos for their tips on molecular biology work and to Dr Lijiang Song and Nicolas for all their help with chemistry. To Dr. Sarah Barry, Lauren, Lona and Dr. Orestis Lazos for proof-reading of my thesis; to Mansoor and David and to all other group members past and present. Big thanks also for our technician Anne for dealing with many practical matters in the lab and for keeping it running on a daily basis.

A very special thank you goes to all the friends I have met while at Warwick, which helped me greatly during the times of my studies, and for their encouragement. A special thank you also goes to my friend Ifti for all the support at the time I needed it.

Finally, I would like to thank the Warwick University and US Department of Agriculture for providing funding for this work.

Declaration

Experimental work contained in this thesis is original research carried out by the author, unless otherwise stated, in the Department of Chemistry at the University of Warwick, between October 2008 and November 2011. No material contained herein has been submitted for any other degree, or at any other institution.

Results from other authors are referenced in the usual manner throughout the text.

_____ Date: _____

Joanna Bicz

Abstract

Streptomyces are Gram-positive bacteria, usually found living within the soil and they are saphrophytes. Among this class of bacteria are some plant pathogenic species, which cause infection of the roots or the tubers of some plants. The model *Streptomycete* plant pathogen is *Streptomyces scabies*; this infects root crops, such as potato or radish and is a known cause of scab disease. Most *Streptomyces* species are producers of secondary metabolites, many of which possess important biological activities, such as antibacterial, iron-chelating, anticancer or immunosuppressant. One group of these secondary metabolites are called siderophores. These are small organic molecules, which can chelate ferric iron. The iron in the environment is mainly present as iron (III) hydroxide, which is not very water soluble and cannot, therefore, be taken up directly by microorganisms. Some bacteria solve this problem through production of siderophores. The siderophores are released into the environment by the microorganisms to chelate iron (III) from the environment and transport it into the cell across the cell membrane. Iron is required for many life processes.
Analysis of the *Streptomyces scabies* genome sequence resulted in the identification of gene clusters predicted to direct the biosynthesis of known siderophores, e.g. desferrioxamines and pyochelin, as well as, potentially novel siderophores. A gene inactivation and comparative metabolic profiling approach has been employed to identify the metabolic products of these gene clusters.

A PCR-targeting method was used to replace part of or whole genes in the *S. scabies* 87.22 putative secondary metabolite gene clusters. An internal fragment of the scabichelin biosynthetic gene *scab85471* and the putative *S. scabies desC* gene were deleted using this method. The scabichelin and *desC* gene mutants were subsequently analysed by LC-MS allowing confirmation of the function of the genes investigated. Production of scabichelin by *S. scabies* 87.22 wild type was analysed by comparing it with the authentic standard. The chemical and genetic complementation of the Δ *desC* mutant was carried out to establish the involvement of the *desC* gene in the biosynthesis of desferrioxamines.

The *S. scabies* 87.22 cryptic tetronate biosynthetic gene cluster predicted to encode a novel agglomerin-like product, which could potentially be involved in plant pathogenicity was also investigated. The expression of the gene cluster was first analysed using reverse transcriptase PCR (RT-PCR) which was carried out on the total RNA isolated from the wild type *S. scabies*. Following this, an attempt was made to disrupt the *scab63021* gene, a putative transcriptional activator of the cryptic tetronate-like cluster in the *S. scabies* genome. Transcriptional analysis of the wild type *S. scabies* and the putative Δ*scab63021* mutant genomes did not show any difference in the expression of the tetronate genes between the wild type strain and the Δ *scab63021* mutant.

Abbreviations

A	Adenylation (domain)
aa	amino acid
ACP	Acyl Carrier Protein
ADP	Adenosine diphosphate
amp	Ampicillin
apra	Apramycin
AT	Acyl Transferase (domain)
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
C	Condensation (domain)
CoA	Coenzyme A
DH	Dehydratase (domain)
DNA	Deoxyribonucleic acid
E	Epimerisation (domain)
EDTA	Ethylenediaminetetraacetic acid
EIC	Extracted ion chromatogram
ER	Enoyl Reductase (domain)
ESI-MS	Electrospray ionization – Mass spectrometry
EtOAc	ethyl acetate
EtOH	Ethanol
FAD	Flavin Adenine Dinucleotide
FRT	FLP recognition targets

GC	Guanine-Cytosine
L-fhOrn	L-N5-formyl-N5-hydroxyornithine
L-hOrn	L-N5-hydroxyornithine
HPLC	High Pressure Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
hyg	Hygromycin
kan	Kanamycin
kb	kilo base pairs
KR	Ketoreductase (domain)
KS	Ketosynthase (domain)
LB	Luria-Bertani (Medium)
LC	(High Pressure) Liquid Chromatography
LC-MS	Liquid Chromatography – Mass Spectrometry
Lys	lysine
M	Methylation (domain)
Me	Methyl
min.	minute
MS	Mass Spectroscopy
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NMR	Nuclear Magnetic Resonance
NRPS	Nonribosomal peptide synthetase
nt	Nucleotide
OBB	Oat Bran Broth (medium)

OD	Optical Density
ORF	Open Reading Frame
oriT	Origin of Transfer
PBGC	pyochelin biosynthetic gene cluster
PCR	Polymerase Chain Reaction
PCP	Peptidyl Carrier Protein
PKS	Polyketide synthase
PLP	Pyridoxal phosphate
PPTase	Phosphopantethienyl transferaze
^{R}	resistance
RBS	Ribosome binding site
rpm	revolutions per minute
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
^{S}	sensitive
SAM	S-adenosylmethionine
Ser	serine
SFM	Soya Flour Mannitol medium
SMM	Supplemented Minimal Medium
T	Thiolation (domain)
Taq	*Thermus aquaticus* (polymerase)
TBE	Tris-boric acid EDTA buffer
TE	Thioesterase (domain)
tet	tetracycline
TO	Time-of-flight

Tris	Tris(hydroxymethyl)aminomethane
UV	ultra-violet

1. Introduction

1.1 *Streptomyces*

Streptomyces belong to the phylum of Actinobacteria and to the family of Streptomycetaceae (Kampfer, 2006). Streptomyces are Gram-positive bacteria, their genomes contain a high content of guanine and cytosine and they have a linear chromosome. They live mainly in soil, where they decompose organic matter to produce their characteristic earthy odour of which the main compound is volatile substance called geosmin (Meyer et al., 1996). Most *Streptomyces* species produce biologically active secondary metabolites and more than two thirds of clinically used antibiotics of microbial origin are produced by *Streptomyces* spp. The biological activities of secondary metabolites also include antimalarial, antifungal and antitumour, as well as some immunosuppressants (Kieser et al., 2000). Some *Streptomyces* spp. are pathogenic, for example, *S. somaliensis* and *S. sudanensis* cause mycetoma in humans (Lichon et al., 2006) ; and *S. acidiscabies*, *S. turgidiscabies* and *S. scabies* cause scab disease in plants (Bignell et al., 2010, 2).

Streptomyces have a complex life cycle (Figure1.1). It begins with the single spore, which germinates to form several multi-nucleoid filaments. These filaments elongate, producing branches, which form hyphae known as vegetative mycelia. After several days the vegetative mycelia grow into air forming aerial hyphae. When the growth of the aerial hyphae is completed it undergoes septation and produces prespore compartments. The spores are assigned into chains and in this step, which is known as a stationary phase, secondary metabolites are produced. The mature spores are then released to start a new life cycle (Kieser et al., 2000).

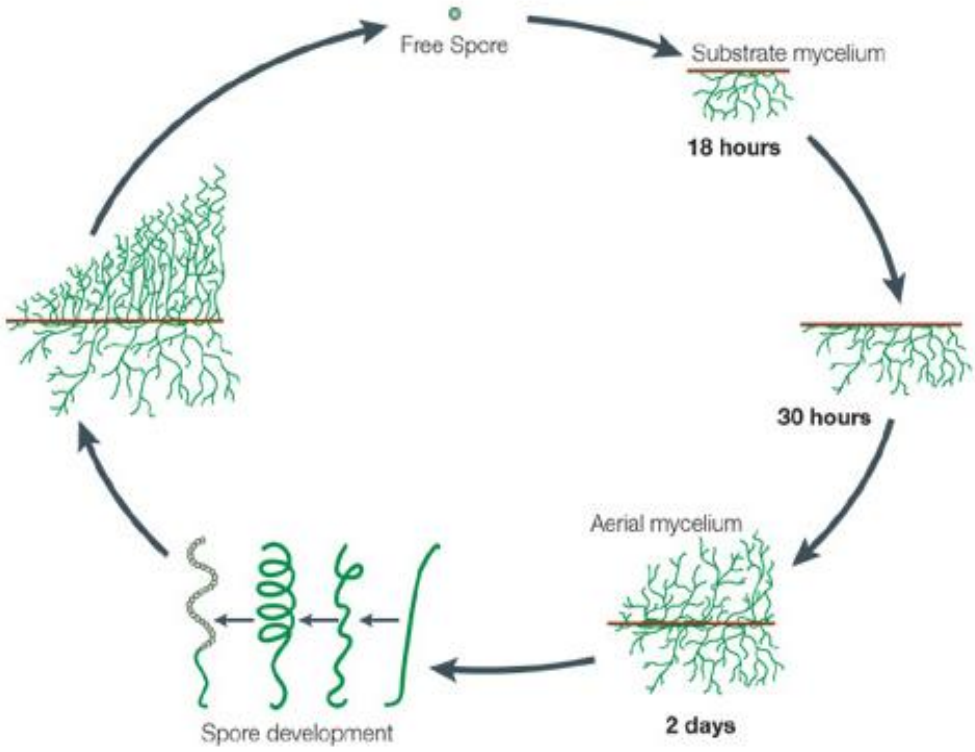


Figure 1.1 The life cycle of *Streptomyces* (Angert, 2005).

1.2 *Streptomyces scabies*

There are over 900 *Streptomyces* species described to date (Labeda et al., 2012), and these include some plant pathogenic species. Three species are *Streptomyces scabies and S. turgidiscabies* and *S. acidiscabies*, which cause infection of root vegetables, such as potato or radish, known as scab disease (Loria and Kempter, 1986; Figure 1.2). *Streptomyces scabies* was one of the first plant pathogens characterised and has a broad world distribution, whereas *S. turgidiscabies* and *S. acidiscabies* species

have been described more recently and have been found only in several isolated world areas (Bignell et al., 2010, 1).

Figure 1.2 Potato and radish with scab disease (http://www.cals.ncsu.edu, http://www.agroatlas.ru).

Some of the secondary metabolites produced by *S. scabies* were known prior to the availability of its genome sequence, including thaxtomins and concanamycins (Kers et al., 2004, Natsume et al., 2005).

One of the first reports of secondary metabolites produced by scab causing streptomycetes was thaxtomins, a family of phytotoxins isolated from *S. scabies* in 1989 (King et al., 1989). It was also shown that *S. scabies* produces other phytotoxins, such as concanamycin A and B (Natsume et al. 1996, Natsume et al. 1998), a macrolide antibiotic first isolated from *Streptomyces diastatochromogenes* S-45 (Kinashi et al., 1984).

1.3 Secondary metabolites of *S. scabies* 87.22

1.3.1 Thaxtomins

The plant pathogenic *Streptomyces* species produce the family of phytotoxins called thaxtomins (Figure 1.3; Loria et al., 2008, Fry and Loria, 2002; Johnson et al., 2009). These toxins, present in infected host tissue (Lawrence et al., 1990), are nitrated dipeptides biosynthesised from L-phenylalanine and L-tryptophan. The 4-nitroindole moiety is required for phytotoxicity (King and Calhoun, 2009). To date, eleven members of the thaxtomin family have been identified, which differ only in the presence or absence of an N-methyl group and the pattern of hydroxylation. Thaxtomin A is the most abundant member of the thaxtomin family (Figure 1.3 B) and is produced by all three plant pathogenic *Streptomyces* species, *S. scabies, S. acidiscabies* and *S. turgidiscabies* (Loria et al., 2008), whereas the main product of *S. ipomeae* is thaxtomin C (King et al., 1994). Thaxtomin A is essential for development of scab disease in plant tissue (King and Calhoun, 2009). The main mode of action of thaxtomins is inhibition of cellulose biosynthesis in growing plant tissue, causing hyperthropy of plant cells, tissue necrosis and cell death (Loria et al., 2008).

Thaxtomin is biosynthesised by non-ribosomal peptide synthetases (NRPSs) encoded by the *txt*A and *txt*B genes (Figure 1.3 A). TxtA and TxtB are responsible for producing the cyclic dipeptide from L-4-nitro-tryptophan and L-phenylalanine. The region close to the *txt*A and *txt*B genes indicated the presence of a gene highly homologous to the oxygenase domain of mammalian nitric oxide synthases (NOSs, Kers et al., 2004). This gene encodes the enzyme TxtD, which provides the nitric

oxide and a cytochrome P450, TxtE, uses it to nitrate L-tryptophan to give the precursor L-4-nitro-tryptophan (Figure 1.3 B; Barry *et. al*., 2012) A cytochrome P450 monooxygenase TxtC is responsible for post-cyclization hydroxylation steps on the thaxtomin precursor to yield thaxtomin A (Healy et al., 2002).

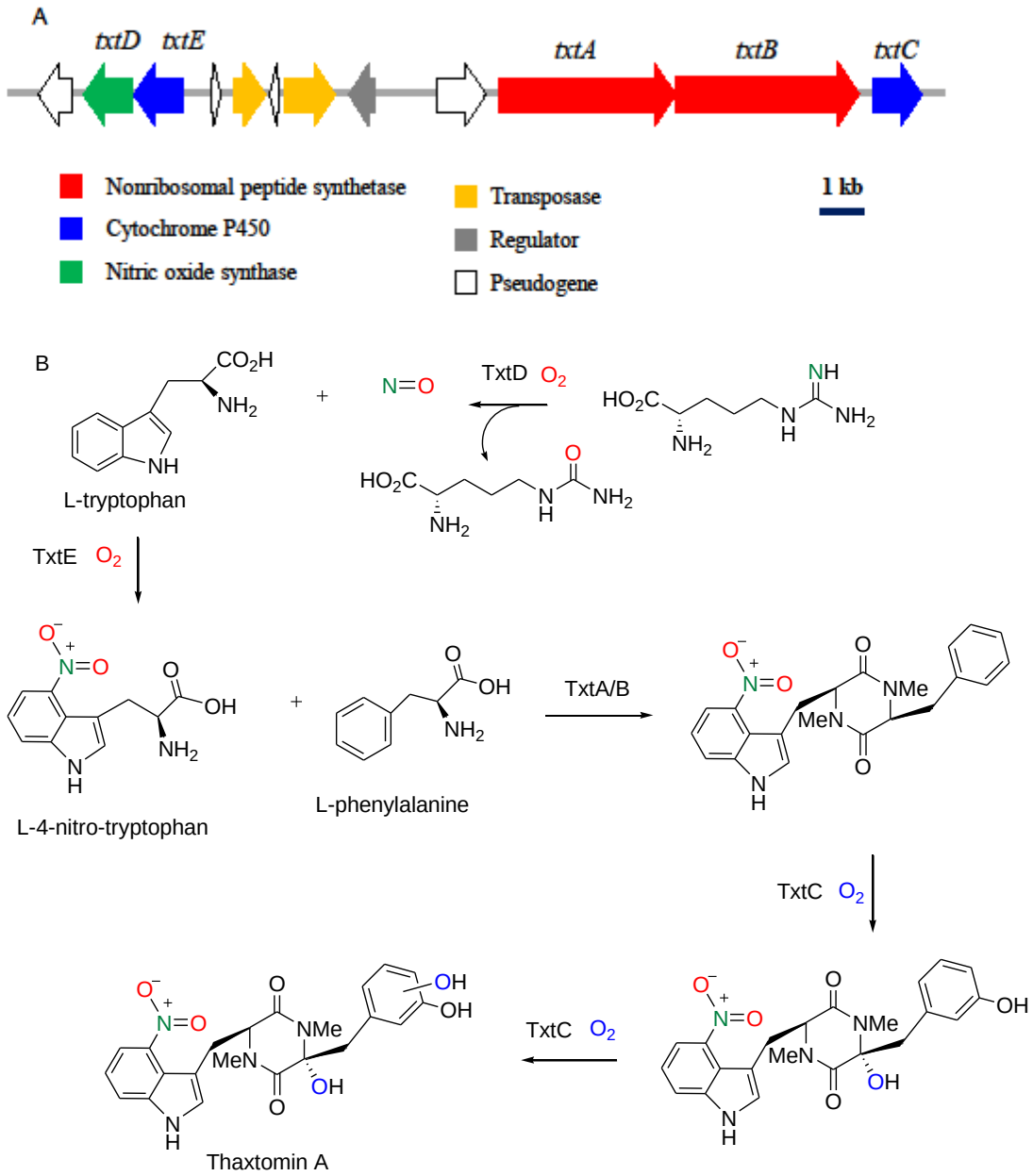


Figure 1.3 A- thaxtomin biosynthetic gene cluster and, B- thaxtomin A biosynthesis (Barry et al, 2012).

1.3.2 Concanamycins

Concanamycins are macrolide antibiotics primarily isolated from *S. diastatochromogenes* (Kinashi et al., 1984), consisting of an 18-membered macrolide ring and a β-hydroxyhemiacetal side chain (Figure 1.4; Haydock et al., 2005). They have a broad range of biological activities, such as antiviral, antiprotozoal, antifungal, and antineoplastic (Kinashi et al., 1984; Omura et al., 1982; Seki-Asano, 1994). Their activity is thought to be due to their inhibition of vacuolar ATP-ases (V-ATP-ases; Drose et al., 1993), the role of which is to actively transport protons into the lumen of various cellular compartments such as endosomes, lysosomes and the vacuoles of plants and fungi (Mellmann et al., 1986; Nelson, 1995). Concanamycins inhibit the acidification of the lysosomes by binding to the V_0 membrane component of the ATP-ase (Drose et al., 2001). Research into these compounds has increased since it was discovered that they may have some application in the treatment of osteoporosis, as V-type ATP-ases play a key role in osteoclast-mediated bone resorption (Laitala-Leinonen et al., 1999).

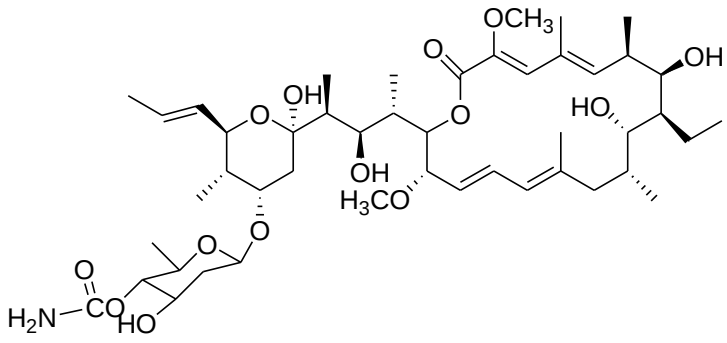


Figure 1.4 Structure of concanamycin A (Haydock et al., 2005).

Concanamycins are produced by a variety of *Streptomyces* species, including *S. diastatochromogenes*, *S. neyagawaensis*, and *S. graminofaciens* (Kinashi et al., 1984; Haydock et al., 2005; Seki-Asano et al., 1994). They have also been found to be produced by *S. scabies* (Natsume et al., 1996) but have not been isolated from *S. acidiscabies* or *S. turgidiscabies* (Natsume et al., 1998). Analysis of the *S. scabies* 87.22 genome indicated that it encodes a putative biosynthetic gene cluster (Figure 1.5 A; Yaxley, 2009) similar to the *S. neyagawaensis* concanamycin biosynthetic cluster (Figure 1.5 B; Nelson, 1995). The *S. scabies* concanamycin biosynthetic cluster contains all the same genes as in the *S. neyagawaensis* concanamycin gene cluster (Yaxley, 2009). The concanamycin cluster in *S. neyagawaensis* is 100 kbp long and encodes 28 genes in total, including six polyketide synthases (PKS's A-F). The 22 other genes present within the biosynthetic gene cluster encode proteins predicted to be involved in sugar biosynthesis, regulatory functions and genes for modification of the polyketide backbone.

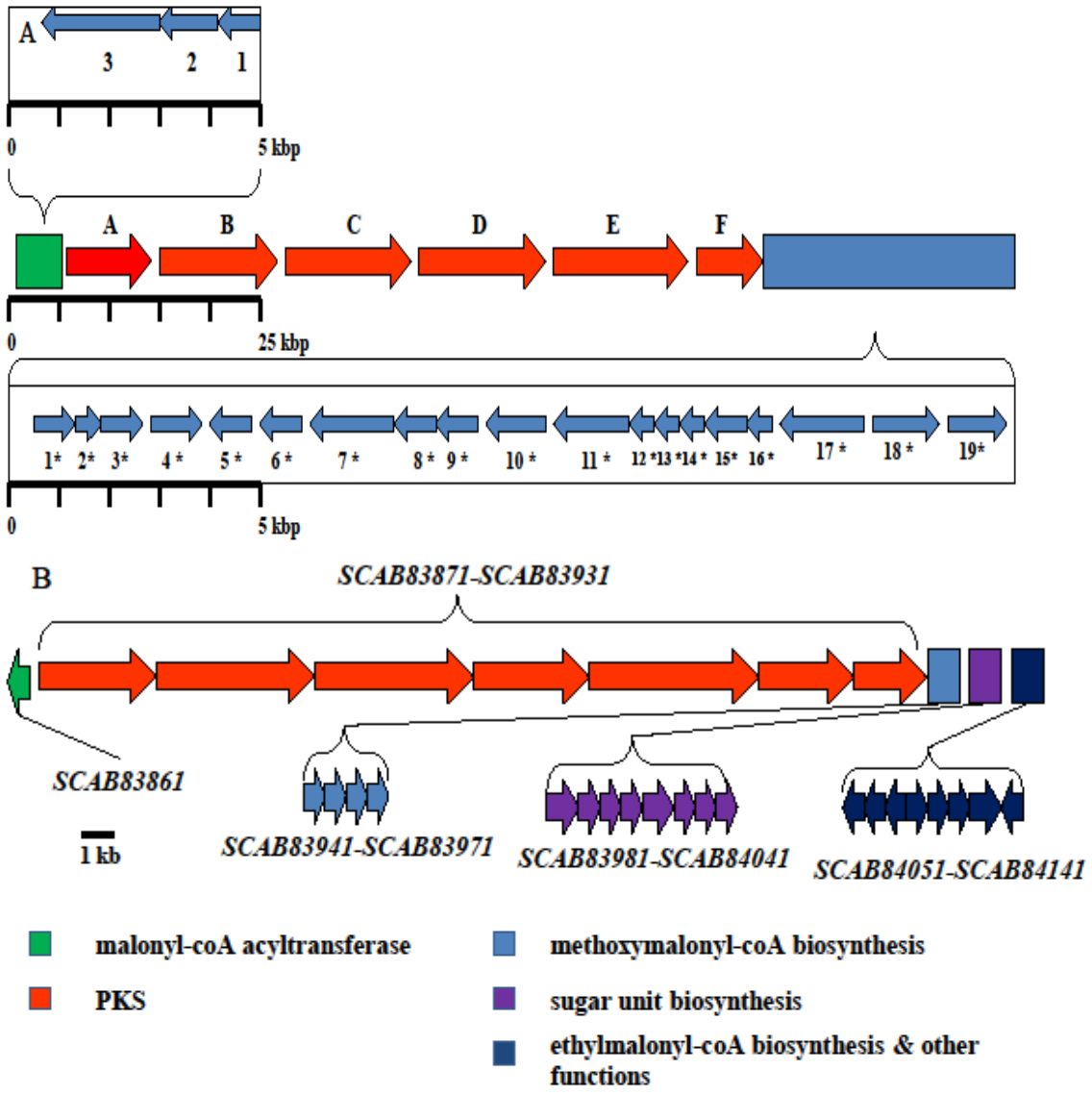


Figure 1.5 Organization of the concanamycin biosynthetic gene cluster in: A- *Streptomyces neyagawaensis* ATCC 27449 (Kinashi et al., 1984). The genes functions are described in the text. B- *S. scabies* 87.22 (Yaxley, 2009).

Concanamycin A contains deoxysugar moiety 4'-O-carbamoyl-2'-deoxyrhamnose, the biosynthesis and attachement of which is encoded for by seven genes ORF5*-ORF11*. ORF5* and ORF6* encode proteins with high homology to a TDP-glucose 4, 6-dehydratase and TDP glucose synthase, respectively (Figure 1.5). ORFs 8*, 9* and 10* encode proteins with high homology to a dTDP-hexose 4-ketoreductase, d-

TDP-sugar-3-ketoreductase and dTDP-sugar-2, 3-dehydratase respectively. ORF7* is highly homologous to carbamoyltransferases and ORF11* is homologous to glycosyltransferases. Assignment of the genes allows predictions to be made with regards to the biosynthetic route towards the deoxysugar moiety. ORF6* is proposed to activate glucose and ORF5* to catalyze 4, 6-dehydration of TDP glucose to form TDP-4-keto-6-deoxyglucose (Figure 1.6). ORF10* is then proposed to catalyse the formation of the 2-deoxy intermediate, which is in equilibrium with the 2, 4-diketo derivative. The ORFs 8*, 9* are then proposed to reduce each keto group at C3 and C4 to give the 2-deoxy derivative with the appropriate stereochemistry at the C3 and C4 positions. ORF7* is proposed to catalyse the addition of a carbamoyl unit to form 4'-O-carbaomyl-2'-deoxyrhamnose. ORF11* is then predicted to catalyse the transfer of the sugar moiety to concanamycin A aglycone (Figure 1.6).

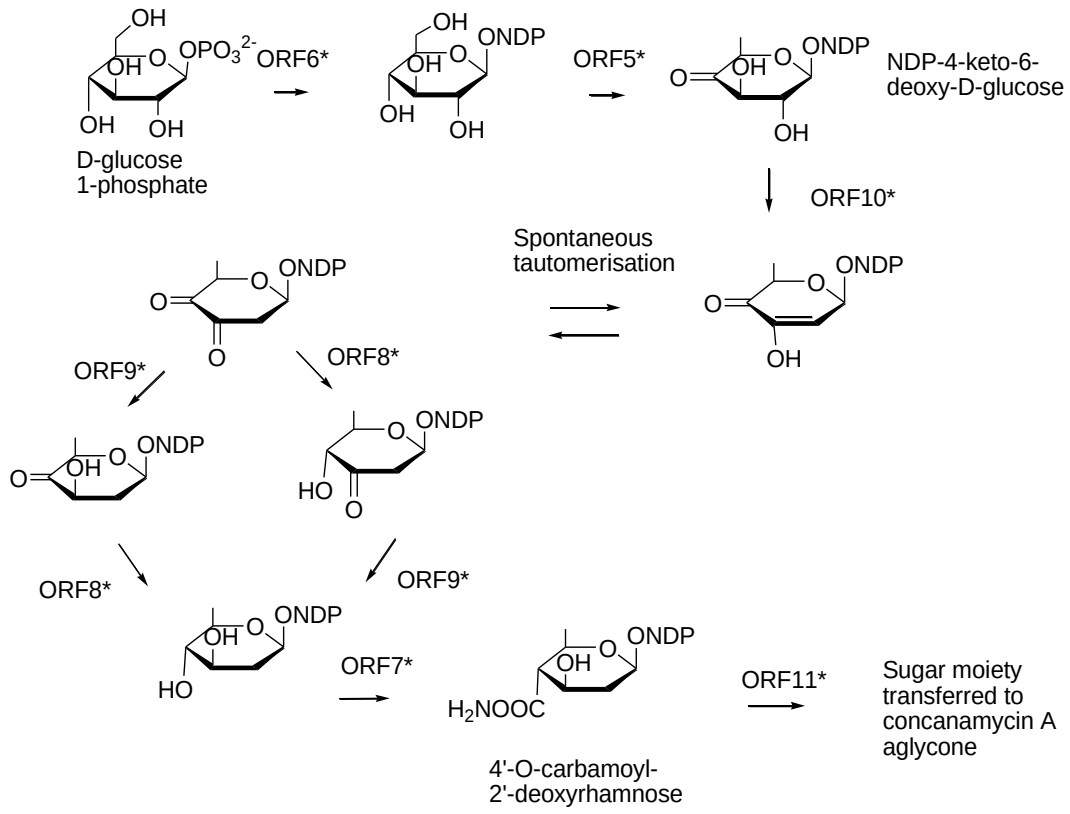


Figure 1.6 Proposed biosynthetic pathway for synthesis and addition of the sugar moiety 4′-*O*-carbamoyl-2′-deoxyrhamnose.

The regulatory functions are proposed to be encoded for by ORF's 2, 3 and 17* (Haydock et al., 2005). ORF3 has high sequence homology to members of the LuxR regulatory family proteins, which are found in the secondary metabolite gene clusters of many Gram-positive bacteria, and ORFs 2 and 17* are similar to the regulatory genes in other PKS gene clusters. ORF17* has 33% identity to AfsR family regulator from *Streptomyces griseus* (Haydock et al., 2005).

ORF1 is highly homologous to malonyl transferase enzymes and its function in the concanamycin A gene cluster is unclear. A putative ORF4B*, located between ORF4* and ORF5*, has no match to any other genes in the databases and its function is unknown. ORFs 18* and 19* encode proteins proposed to be involved in

biosynthesis of extender units. The enzymes are predicted to be a crotonyl-CoA reductase and a β-hydroxybutyryl-CoA reductase, respectively, which could both play a role in the biosynthesis of a butyrate extender unit for module 10. The extender unit selected for by modules 6 and 13 of the PKS could be a result of direct incorporation of methoxymalonyl-ACP. ORF12*, is homologous to FkbG and may be involved in the biosynthesis of methoxymalonyl-ACP. ORF1* is homologous to FkbB, ORF2* to FkbJ, ORF3* to FkbI and ORF4* to FkbH. These proteins are predicted to catalyse the conversion of primary metabolites into methoxymalonyl-ACP.

The biosynthesis of the concanamycin A polyketide backbone in *S. neyagawaensis* is directed by the operon of six genes *conA* to *conF* (Figure 1.7), which encode the PKS enzymes CON1-6, respectively. CON1 (2841 aa, approx. 300 kDa) contains the loading module and module 1 for chain extension. CON2 (4825 aa, approx. 500 kDa) contains extension modules 2-4. CON3 (4970 aa, approx. 520 kDa) contains extension modules 5–7, CON4 (4490 aa, approx. 430 kDa) contains modules 8–9 and CON5 (5599 aa, approx.530 kDa) contains modules 10–12. CON6 (2056 aa, approx. 215 kDa) contains the final condensation module 13 and a C-terminal thioesterase domain for cyclization and release of the polyketide chain (Figure 1.7; Haydock et al., 2005). More about PKSs family proteins is discussed in detail in section 1.8.

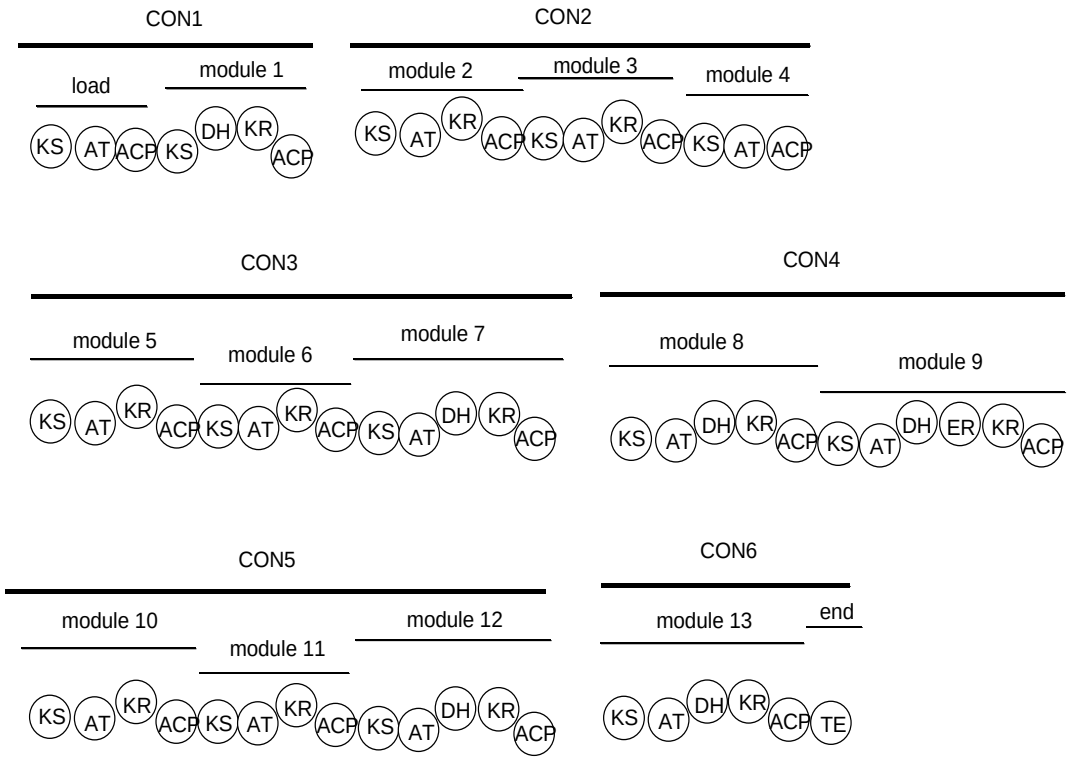


Figure 1.7 Domain organisation of concanamycin A PKS genes (Haydock et al., 2005).

1.4 Secondary Metabolites of *S. scabies* 87.22 Discovered by Genome Mining.

S. scabies 87.22 is a model organism for studying plant pathogenicity in Gram-positive bacteria (Lambert and Loria, 1989). The sequence of the genome of *S. scabies* became available through the *S. scabies* genome-sequencing project (www.sanger.ac.uk/Projects/S.scabies). The genome is 10.1 Mb long with a GC content of 72%. The availability of the genome sequence for *S. scabies* allowed comparative genomic analyses with other bacterial genomes, to identify several biosynthetic gene clusters proposed to encode either known (such as

desferrioxamines, pyochelin) or novel (agglomerin-like product, coronofacic acid-like product) secondary metabolites (http://www.ncbi.nlm.nih.gov/nuccore/260644157, http://strepdb.streptomyces.org.uk).

Some of the biosynthetic gene clusters identified were found to be conserved in the genomes of other *Streptomyces* species (e.g. *des* gene cluster in *S. coelicolor* A(3) 2) and other sequenced microbes (e.g. the pyochelin biosynthetic gene cluster in *P. aeruginosa*; Table 1.1; Yaxley, 2009). The products of some of secondary metabolite biosynthetic gene clusters are currently unknown; such clusters are referred to as 'cryptic' biosynthetic gene clusters.

Table 1.1 Overview of *S. scabies* known and putative natural product biosynthetic gene clusters (Yaxley, 2009). Yellow- biosynthetic gene clusters conserved both in *S. coelicolor* and *S. avermitilis*; white-biosynthetic gene clusters conserved either in *S. coelicolor* or in *S. avermitilis*; blue- clusters conserved in neither of the two species.

Biosynthetic system	**Metabolite**	***S. scabies* gene cluster**	**Estimated size/kb**	**Conserved in *S. coelicolor*/*S. avermitilis***
Bacterial type II PKS	spore pigment ?	*scab43271 - scab43341*	8	both
NRPS-independent siderophore synthetase	siderophore ?	*scab24661- scab24751*	16	both
NRPS-independent siderophore synthetase	desferrioxamines ?	*scab57921- scab57981*	9	both
NRPS-independent siderophore synthetase	siderophore ?	*scab18371- scab18421*	7	both

Ectoine synthase	ectoine compatible solute	*scab70711-scab70751*	4	both
Hopene/squalene synthase	pentacyclic hopanoids	*scab12951-scab13061*	14	both
Phytoene/polyprenyl synthase	carotenoid pigment	*scab5431-scab5511*	12	both
Sesquiterpene synthase	geosmin?	*scab20121*	2	both
Tyrosinase	melanin pigment	*scab59231-scab59241*	2	both
Tyrosinase	melanin pigment	*scab85681-scab85691*	2	both
Bacterial type III polyketide synthase	germicidins?	*scab80171*	4	in *S. coelicolor*
NRPS-independent siderophore synthetase	siderophore ?	*scab84501-scab84521*	8	in *S. avermitilis*
NRPS	unknown	*scab72991*	4	in *S. avermitilis*
Polysaccharide	exopolysaccharide	*scab23341-scab23551*	29	in *S. coelicolor*
NRPS	pyochelin siderophore ?	scab1411-scab1571	35	no
NRPS	NRPS siderophore?	scab85431-scab85521	42	no
NRPS	lipopeptide?	scab19681-scab19731	34	no
NRPS	unknown	scab19681-scab19731	8	no
NRPS	thaxtomin	scab31761-scab31841	19	no
Hybrid NRPS/PKS system	unknown	scab62901-scab63011	14	no
Hybrid NRPS/PKS system	unknown	scab78961-scab78981	8	no

Hybrid NRPS/PKS system	unknown	scab43961	13	no
Bacterial type I PKS	macrolide phytotoxin ?	scab83871-scab84091	95	no
Bacterial type I PKS	coronofacic acid derivative ?	scab79581-scab79721	31	no
Class II DAHP synthase/AfsA homologue	butyrolactone ?	scab12021-scab12111	12	no
Lantibiotic	lantibiotic ?	scab31961-scab32051	14	no
Mixed	unknown	scab63251-scab63401	8	no
Mixed	unknown	scab69771-scab69871	13	no
Other	bothromycin	scab56711-scab56591	18	no

1.4.1 Coronofacic acid

Analysis of the *S. scabies* 87.22 genome sequence revealed the presence of a secondary metabolite gene cluster that is predicted to encode a compound homologous to coronofacic acid (CFA). CFA forms part of coronatine (COR), a phytotoxin produced by some plant pathogenic bacteria such as *Pseudomonas syringae* (Figure 1.8 A; Bignell et al., 2010, 2). In this strain COR is an important virulence factor and the *cfa* biosynthetic cluster is conserved in that strain as well as in the plant pathogen, *Pectobacterium atrosepticum*. Analysis of genomes of nonpathogenic *Streptomyces* (*S. coelicolor*, *S. avermitilis*, *S. griseus*) showed that they do not encode the *cfa* biosynthetic cluster. As *cfa* biosynthetic cluster is conserved amongst some plant pathogenic bacteria, it is possible that the *S. scabies* *cfa* cluster may also be important for virulence in this strain.

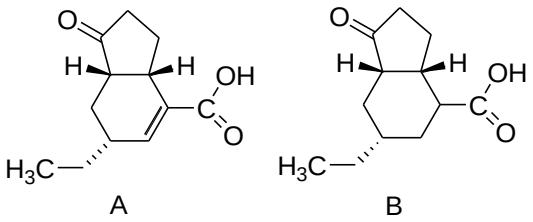


Figure 1.8 A-Structure of coronafacic acid (CFA) from *Pseudomonas syringae* and, B- A predicted structure for the CFA-like compound produced by *S. scabies* 87.22 (Bignell et al., 2010, 2).

The *cfa* biosynthetic cluster in *S. scabies* is 31 kb long and includes 15 genes (Figure 1.9; Bignell et al., 2010, 2), nine of which are homologous to those in the *Pseudomonas syringae* and the *Pectobacterium atrosepticum cfa* clusters.

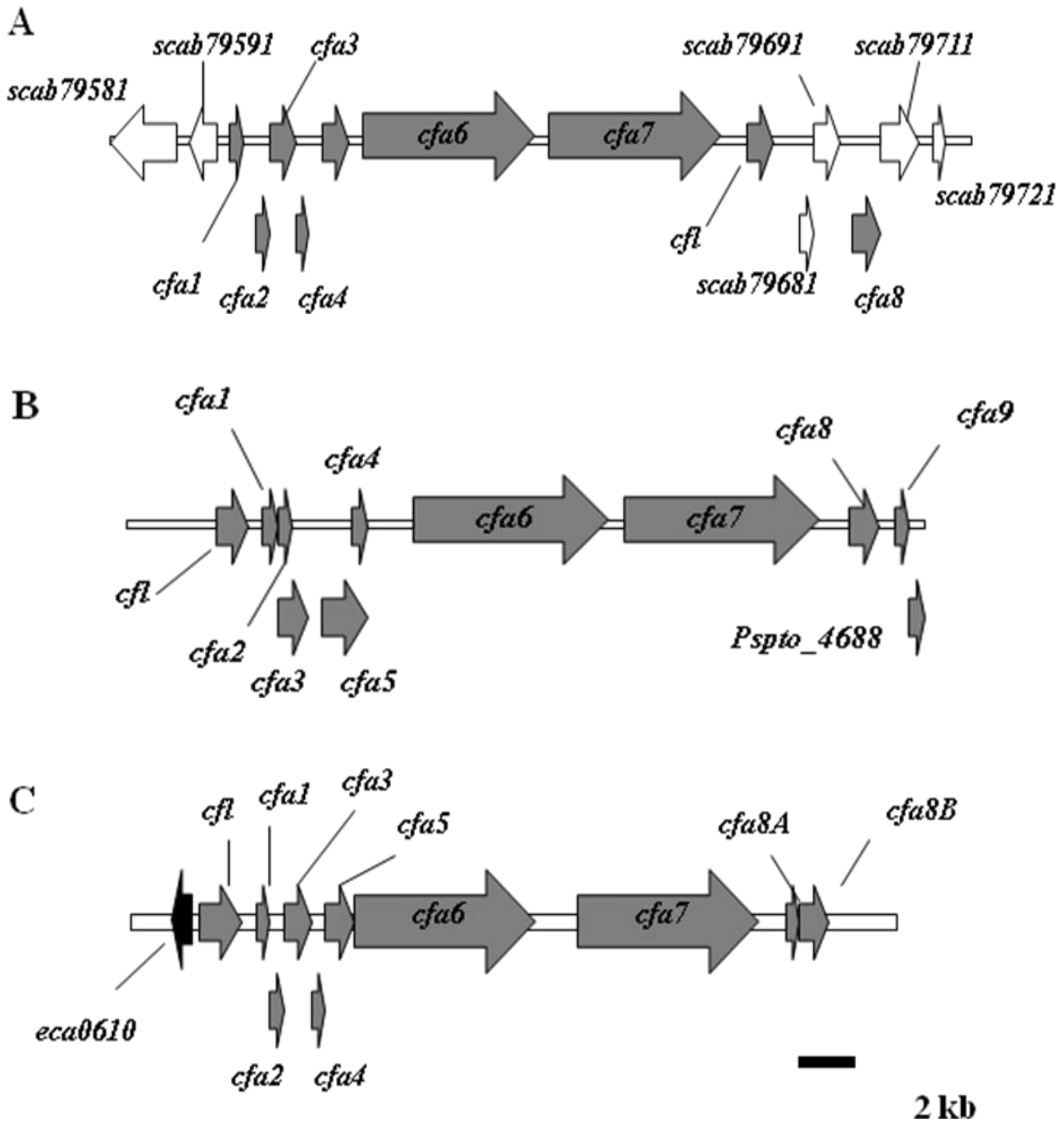


Figure1.9 A-The *cfa*-like biosynthetic gene cluster of *Streptomyces scabies* 87.22; B-the *cfa* biosynthetic cluster of *Pseudomonas syringae* pv. tomato DC3000; and C-the putative *cfa* biosynthetic cluster of *Pectobacterium atrosepticum* SCRI1043 (Bignell et al., 2010, 2).

The genes *cfa1-5* encode enzymes predicted to catalyse the synthesis of a CFA intermediate 2-carboxy-2-cyclopentenone (CPC; Rangaswamy et al., 1998). The genes *cfa6* and *cfa7* encode type-I PKSs catalysing the biosynthesis of the CFA backbone. The *S. scabies* Cfa6 and Cfa7 enzymes are similar to their *Pseudomonas*

syringae and *Pectobacterium atrosepticum* analogs, however, the enoyl reductase (ER2) domain in the *S. scabies* Cfa7 PKS is not present in the *Pseudomonas syringae* and *Pectobacterium atrosepticum* equivalent PKS (Bignell et al., 2010, 2).

One more gene in the *S. scabies cfa* cluster that is homologous to the genes in the *Pseudomonas syringae* and *Pectobacterium atrosepticum* cluster is coronofactate ligase (*cfl*). Its function in *Pseudomonas syringae* is predicted to be ligation of CFA to coronamic acid (CMA) to produce coronatine (Figure 1.10). Although *cfl* is present in *S. scabies* and *Pectobacterium atrosepticum* these strains are not thought to produce coronatine as they do not encode genes for CMA biosynthesis (Bell et al., 2004).

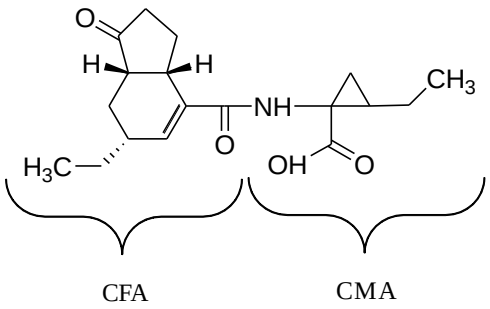


Figure 1.10 Structure of the *Pseudomonas syringae* coronatine (COR) phytotoxin consisting of coronafacic acid (CFA) linked to coronamic acid (CMA) (Bignell et al., 2010, 2).

The six other genes that are encoded in the *S. scabies* biosynthetic cluster and which have no homologues in *Pseudomonas syringae* or *Pectobacterium atrosepticum* clusters are: *scab79581*, *scab79591*, *scab79681*, *scab79691*, *scab79711*, *scab79721*. *Scab79581* encodes a ~640 amino acid protein the function of which is unknown. It has low homology to other proteins from the BlastP database. It is 25% identical to a

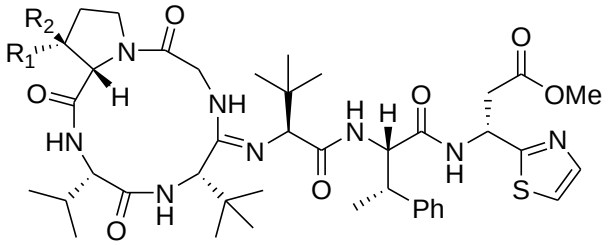
UBA/ThiF-type NAD/FAD-binding fold protein from *Nitrosospira multiformis* (Bignell et al., 2010, 2). *Scab79591* encodes a 266 amino acid protein with highest similarity to transcriptional regulators from *Streptomyces* antibiotic biosynthetic clusters (LuxR family). The genes *scab79681* and *scab79721* encode a putative oxidoreductase and a dehydrogenase/reductase, respectively. They are thought to be involved in the modification of the polyketide compound produced by *S. scabies*. *Scab79691* encodes a protein with sequence similarity to P450 monooxygenases that are typically involved in introducing hydroxyl groups in synthesized polyketide products. *Scab79711* encodes a putative 3-hydroxybutyryl-CoA dehydrogenase, which may have a role in the reduction of acetoacetyl-CoA to 3-hydroxybutyryl-CoA, an intermediate in crotonyl-CoA biosynthesis. Crotonyl-CoA is a substrate for the crotonyl-CoA carboxylase/reductase (CCR), which converts Crotonyl-CoA into an ethylmalony-CoA extender unit (Chan et al., 2009).

Analysis of the genes from the *S. scabies* cfa-like biosynthetic cluster suggests that *S. scabies* may produce a compound, which has similar structure to the coronofacic acid from *P. syringae* (Figure 1.8 B). Transcriptional studies indicated that the cfa-like cluster is expressed in *S. scabies* in thaxtomin inducing media (Bignell et al., 2010, 2).

1.4.2 Bottromycin

Bottromycin (Figure 1.11) is a macrocyclic peptide antibiotic that was originally isolated from *Streptomyces bottropensis* in 1957 (Waisvisz et al., 1957). It exhibits activity against Gram-positive bacteria and mycoplasma, including methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococci*

(VRE; Kobayashi et al., 2010). Its antibacterial properties result from inhibiting bacterial protein synthesis by blocking an aminoacyl-tRNA binding to the A site of bacterial ribosomes (Otaka and Kaji., 1976; Otaka and Kaji., 1981; Otaka and Kaji., 1983). It is comprised of a macrolactamidine, non-proteinogenic amino acids and a thiazole ring and its total synthesis was elucidated in 2009 (Shimamura et al., 2009).



bottromycin A2 (R1 = Me, R2 = H)
bottromycin B2 (R1 = R2 = H)
bottromycin C2 (R1 = R2 = Me)

Figure 1.11 Structure of bottromycin (Gomez-Escribano et al., 2012).

Using a combination of high-throughput DNA sequencing, genetic manipulation and comparative metabolite profiling the bottromycin biosynthetic gene cluster was identified in *S. bottropensis* (Figure 1.12 A; Gomez-Escribano et al., 2012). Bioinformatic analysis of bottromycin biosynthetic gene cluster unveiled an almost identical gene cluster within the genome of *S. scabies* 87.22 (Figure 1.12 B; Crone et al., 2012). *S. scabies* had not been reported to produce bottromycin, consequently it was analysed for bottromycin production alongside *S. bottropensis* using LC-MS (Figure 1.13). This demonstrated that bottromycins are also produced by *S. scabies* (Gomez-Escribano et al., 2012; Crone et al., 2012).

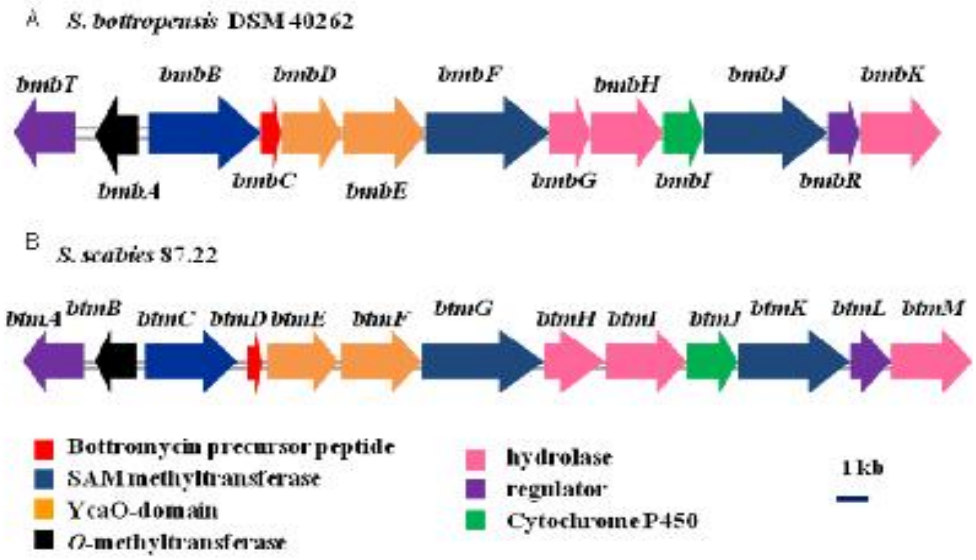


Figure 1.12 Putative bottromycin biosynthetic gene clusters in *S. bottropensis* DSM 40262 (Gomez-Escribano et al., 2012) and *S. scabies* 87.22 (Crone et al., 2012).

Analysis of the *S. bottropensis* genome identified eleven putative biosynthetic genes (*bmbA-K*, Figure 1.12A), as well as, gene encoding a putative pathway specific regulatory protein (*bmbR*) and a gene encoding a putative exporter of the assembled product (*bmbT*). The gene cluster appears to be organised into two operons: *bmbA-T* and *bmbB-K*. The putative bottromycin biosynthetic gene cluster in *S. scabies* comprises of the genes *scab56711* to *scab56591*, which were denoted *btmA-M* (Figure 1.12 B). The functions of bottromycin biosynthetic genes from *S. bottropensis*, as well as, their *S. scabies* homologs are listed in Table 1.2.

Table 1.2 Genes and proteins involved in bottromycin biosynthesis and their putative functions (adapted from Gomez-Escribano et al., 2012).

Gene name	***S. scabies* homolog**	**Motifs**	**Putative function**
bmbT	Scab56711 (*btmA*)	putative integral membrane protein	Bottromycin export and immunity
bmbA	Scab56701 (*btmB*)	Leucine carboxyl methyltransferase	O-methylation of Asp residue
bmbB	Scab56691 (*btmC*)	B12 binding/Radical SAM	β-methylation of Phe/Val/Pro residues
bmbC	Scab56681 (*btmD*)	Contains core peptide sequence GPVVVFDC	Precursor peptide
bmbD	Scab56671 (*btmE*)	YcaO-like family	Thiazoline formation?
bmbE	Scab56661 (*btmF*)	YcaO-like family (shows similarity to microcin B17 thiazoline/oxazoline synthetase McbD)	Thiazoline formation
bmbF	Scab56651 (*btmG*)	B12 binding/Radical SAM	*O*-methylation of Phe/Val/Pro residues
bmbG	Scab56641 (*btmH*)	Abhydrolase 1 family, alpha/beta hydrolase fold domain	Peptidase involved in macrolactamidine formation
bmbH	Scab56631 (*btmI*)	Metal-dependent hydrolase	Peptidase involved in macrolactamidine formation
bmbI	Scab56621 (*btmJ*)	Cytochrome P450	Oxidative decarboxylation of thiazoline to form thiazole
bmbJ	Scab56611 (*btmK*)	B12 binding/Radical SAM	β-methylation of Phe/Val/Pro residues
bmbR	Scab56601 (*btmL*)	DNA-binding domain of ArsR; DUF2087	Putative transcriptional regulator
bmbK	Scab56591 (*btmM*)	Peptidase M17 cytosolic aminopeptidase family	Removal of N-terminal Met residue

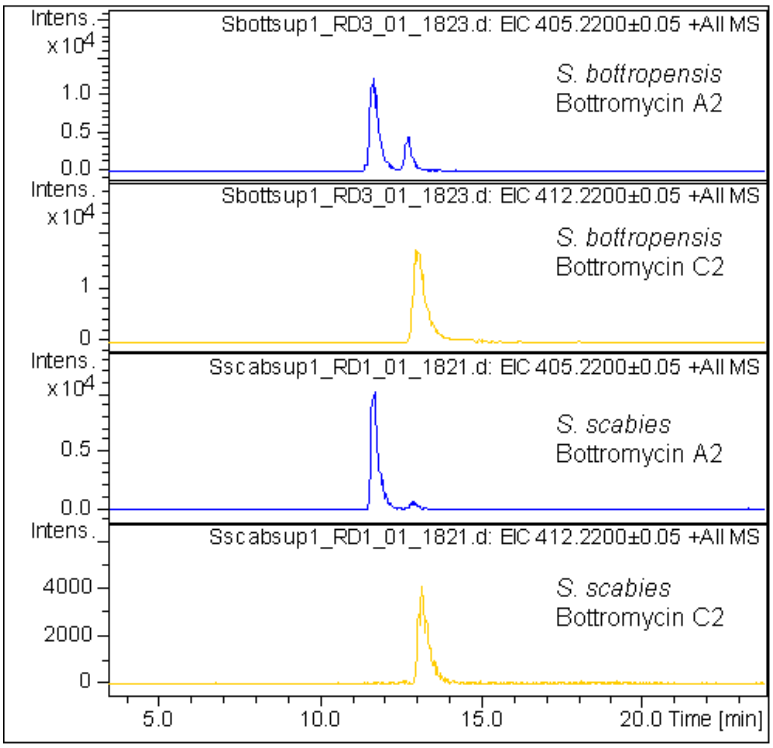


Figure 1.13 EIC for *m/z* = 405.2200 and 412.2300 (corresponding to the $[M+2H]^{2+}$ ions for bottromycins A2 and C2, respectively) from LC-MS analyses of culture supernatants of *S. bottropensis* DSM40262 (top two traces) and EIC for *m/z* = 405.2228 and 412.2314 (corresponding to the $[M+2H]^{2+}$ ions for bottromycins A2 and C2, respectively) from LC-MS analyses of culture supernatants of *S. scabies* (bottom two traces; Gomez-Escribano et al., 2012).

Labelled precursor feeding experiments (Kellenberger, 1997) lead to the hypothesis that bottromycins are biosynthesised from a ribosomal precursor peptide containing the sequence GPVVVFDC. This undergoes several unusual posttranslational reactions, such as β-methylation of the Phe, Pro and two Val residues, O-methylation of the Asp residue, proteolytic lactamidine formation, and conversion of a thiazoline to a thiazole via hydroxylation of the Cys β-carbon followed by decarboxylation coupled with dehydration .

Three putative methyl-cobalamin (MeCob)-dependent radical S-adenosylmethione (SAM) enzymes are encoded for by *bmbB, bmbF* and *bmbJ*. These are proposed to

catalyse β-methylation of the Phe, Pro and two of the three Val residues in BmbC (Figure 1.14). The putative SAM-dependent O-methyltransferase, encoded for by *bmbB*, is proposed to convert the carboxylic acid group of the Asp residue to the corresponding methyl ester (Figure 1.14). The protein similar to oxazoline/thiazoline synthetases (encoded for *bmbC* gene) is involved in the biosynthesis of many other heterocycle-containing ribosomal peptides (Dunbar et al., 2012). It is hypothesised that BmbE forms a complex with BmbD, similar to oxazoline/thiazoline synthetases, which then catalyses ATP-dependent thiazoline formation. Proteases thought to be encoded for by *bmbG*, *bmbH* and *bmbK* are proposed to catalyse the removal of the N-terminal N-formyl-Met residue and the 35 amino acids located downstream of the Cys residue. Removal of the N-terminal N-formyl-Met residue enables nucleophlic attack of the Gly amino group on the carbonyl group of the appropriate β-methyl-Val residue with subsequent collapse of the resulting intermediate with loss of water, which results in macrolactamidine formation. This reaction is proposed to be catalysed by the third protease. The putative cytochrome P450 is encoded for by *bmbI* and it is thought to catalyse the conversion of the thiazoline to a thiazole via hydroxylation of the β-carbon of the Cys residue, followed by elimination of water and subsequent decarboxylation. The final reaction in the biosynthetic pathway is proposed to be the epimerisation of the Asp β-carbon, yielding bottromycin A2. It is not known whether this reaction is catalysed by a specific enzyme or is spontaneous. Bottromycins B2 and C2 arise from the same pathway, except that C-methylation of the Pro residue is skipped in the biosynthesis of the former and occurs twice in the assembly of the latter. *bmbT* encodes a putative integral membrane protein that may be involved in the export of bottromycin. *bmbR* is proposed to encode a negative

transcriptional regulator of the bottromycin biosynthetic gene cluster, as it has 32% identity to ArsR of *Roseiflexus castenholzii* (Gomez-Escribano et al., 2012).

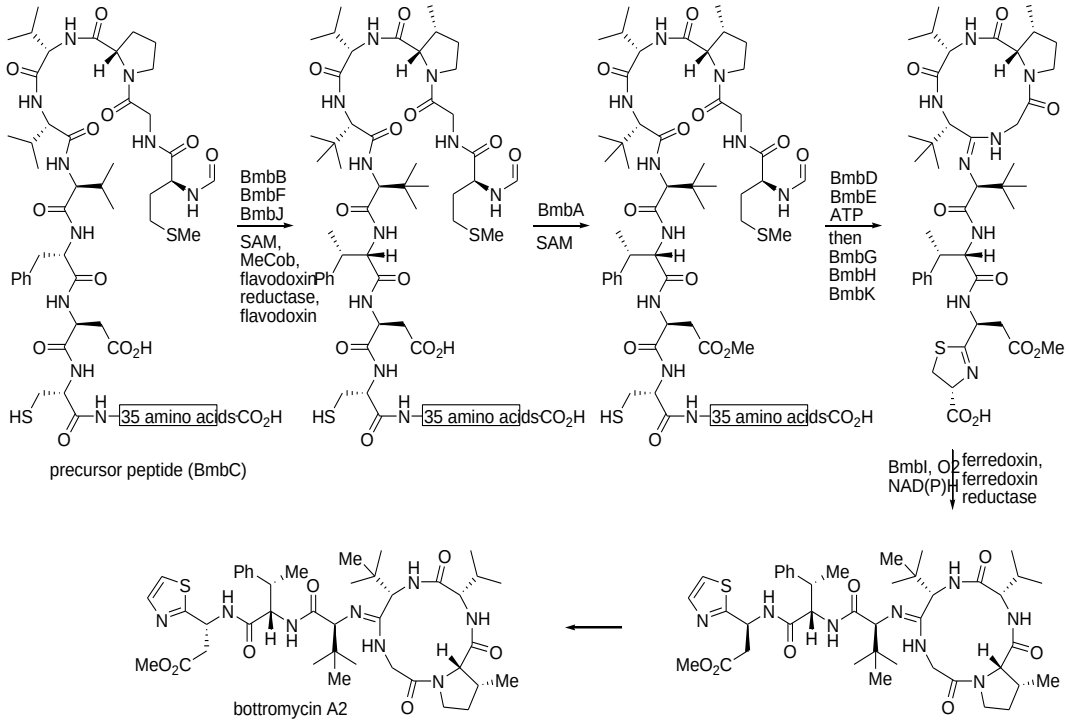
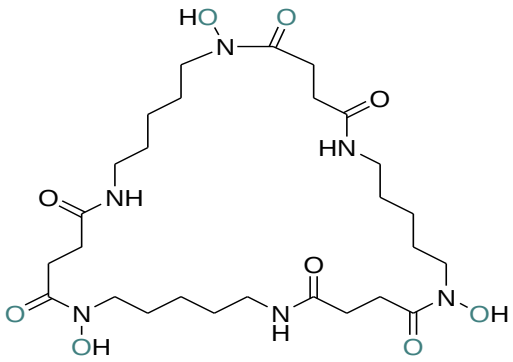


Figure 1.14 Proposed pathway for bottromycin biosynthesis in *S. bottropensis* and *S. scabies* (Gomez-Escribano et al., 2012).

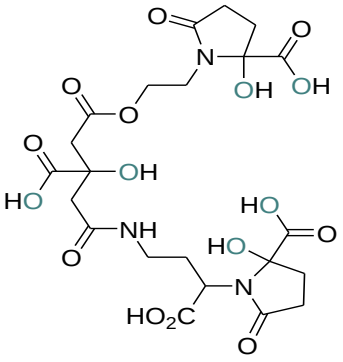
1.5 Siderophores

Siderophores are low molecular weight organic molecules produced by bacteria and fungi (Neilands, 1995). The iron is present in the environment mainly in the ferric form, for example as iron (III) hydroxide, which is not very soluble in water at neutral pH; therefore, it cannot be taken up directly by microorganisms.To overcome this many bacteria secrete siderophores into the environment, where they strongly chelate iron (III). The iron complex is then actively transported back into the cell (Kraemer, 2005).

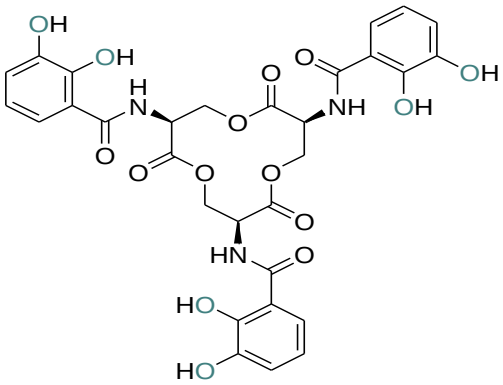
The ability of the siderophore to chelate iron results from formation of stable multidentate (e.g. hexadentate) complexes with iron (III). This is possible due to the presence of iron-chelating ligands in the siderophores structure. The most commonly observed are one to several bidentate oxygen ligands, such as hydroxamates (e.g. in desferrioxamine E), catecholates (e.g. in enterobactin), or α-hydroxycarboxylates (e.g. in achromobactin; Figure 1.15; Miethke and Marahiel, 2007). Some siderophores contain oxygen and nitrogen heterocyclic structures to chelate ferric iron (e.g. pyochelin; Figure 1.15; Miethke and Marahiel, 2007).



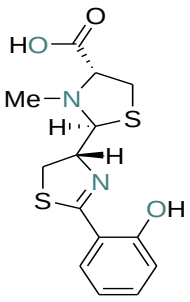
desferrioxamine E



achromobactin



enterobactin



pyochelin

Figure 1.15 Structures of several siderophores (chelating atoms are highlighted in blue; Miethke and Marahiel, 2007).

Iron is required for many mammalian cell processes, such as gene regulation, DNA synthesis, oxygen transport, respiration and lipid metabolism to name a few (Recalcati et al., 2006). To perform these functions iron is incorporated into haemoglobin, myoglobin and cytochromes. It is also a part of non-haeme structures, e.g.iron-sulfur motifs found in the prosthetic groups of some proteins (Meyer, 2008). In plants, iron is required for chlorophyll formation and for metabolism (Marsh et al., 1963). Iron is also essential for bacterial growth. Iron acquisition, through the production and release of siderophores, is necessary for the succesful infection of mammalian hosts by bacterial pathogens (Miethke and Marahiel, 2007). For example, mycobactin, a siderophore produced by *M. tuberculosis*, is able to extract iron from macrophages (Luo et al., 2005). Siderophore-mediated iron acquisition can be inhibited by the mammalian host protein siderocalin, which binds to siderophores released by the pathogen. However, some bacteria have developed a way to circumvent this defense mechanism through the production of siderophores, to which, siderocalin cannot bind. An example of this resistance mechanism is found in *Vibrio cholera*, which produces the siderophore vibriobactin (Li et al., 2012).

1.5.1 Siderophores Biosynthesised via NRPS-Dependent Pathways

Many siderophores are biosynthesised by enzymes belonging to the family of nonribosomal peptide synthetases (NRPSs; Challis, 2005). They catalyse the condensation of amino acids to form biologically-active peptides. These molecules commonly possess structural features, such as non-proteinogenic amino acids, hydroxy acids, heterocyclic rings and other unusual modifications in their backbone. Products of NRPSs include not only siderophores but other pharmacologically

significant compounds, for example the antibiotics penicillin (Smith et al., 1990), vancomycin (Wageningen et al., 1998), gramicidin (Kessler et al., 2004) and immunosuppressive agents such as cyclosporine (Finking and Marahiel, 2004).

NRPS enzymes consist of a number of repeat units referred to as modules. Each module is generally responsible for the extension of the growing peptide chain by one amino acid residue, as well as any modifications to be made to the selected amino acid. The number and order of amino acids in the peptide being synthesized corresponds to the order of modules in the NRPS, with exceptions, such as coelichelin (Barona-Gomez, 2006).

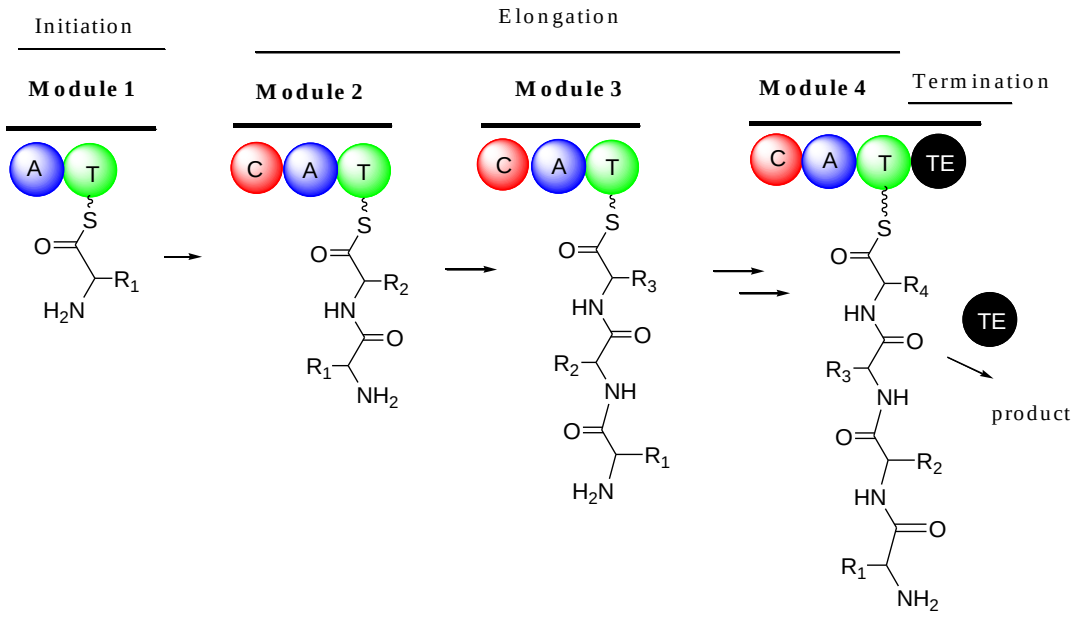


Figure 1.16 Steps of nonribosomal peptide biosynthesis, domains: A – adenylation, T (PCP) – thiolation, C – condensation, TE – thioesterase.

Each module consists of a set of different domains. Each NRPS chain elongation modules usually contains adenylation (A), thiolation (T) (also known as the peptidyl carrier protein-PCP), and condensation (C) domains (Figure 1.16; Challis and Lautru, 2004). These chain elongation modules may also contain additional or

alternative domains such as epimerisation (E), methylation (M), and heterocyclization (Cy), which introduce modifications to the amino acid incorporated by the module into the biosynthesized polypeptide chain. Another two modules usually present in the NRPS are a loading module, consisting only of A and T domains; and a chain terminating module containing the thioesterase (TE) domain (Figure 1.16).

The adenylation (A) domain is considered to have a primary role in substrate selectivity of the module. During nonribosomal peptide biosynthesis (Figure 1.16), the A domain firstly recognises the amino acid being incorporated, activates it as an amino acyl adenylate and then transfers it to the adjoining T domain (Figure 1.17). The activated amino acid is linked to the T domain via a covalent bond as a thioester to the 4'-phosphopantetheinyl (4'Ppant) arm of the T domain. The 4'Ppant is posttranslationally attached to a conserved serine of the T domain. This reaction is carried out by a broad specificity phosphopantetheinyl transferase, which catalyses the attack of a conserved serine residue in the T domain on the diphosphate group of coenzyme A. 3', 5'-adenosine diphosphate (ADP) is released and the T domain with a phosphopantetheine thiol group is generated (Figure 1.18; Walsh et al., 1997).

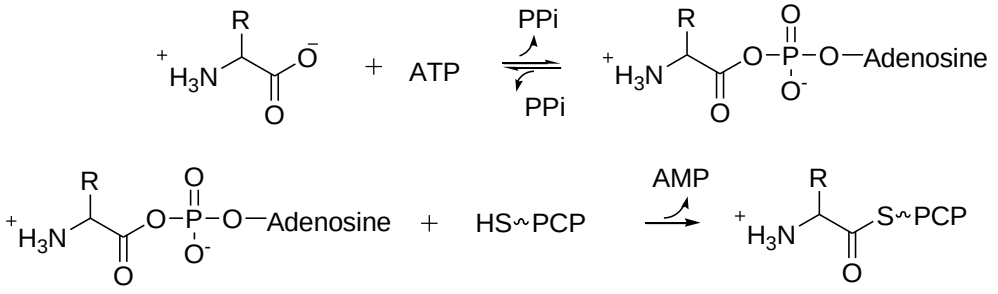


Figure 1.17 The A domain reaction catalysis: the activation of the amino acid substrate through the formation of an aminoacyl adenylate followed by the covalent binding via a thioester bond to the 4'Ppant cofactor of the PCPdomain.

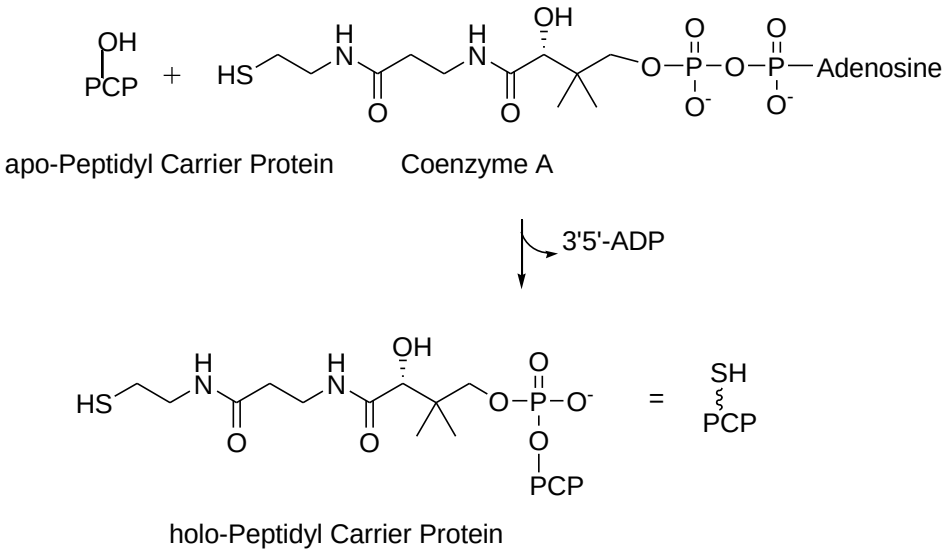


Figure 1.18 Reaction catalysed by PPTases. The *apo* is the inactive form, lacking the phosphopantetheinyl arm and the *holo* is the active form.

Elongation of the growing peptide chain is then carried out by the condensation (C) domain, which joins the amino acid attached to the T domain of its module with an amino acid attached to the T domain of the downstream module. This is achieved by the nucleophilic attack of α-amino group of the peptidyl chain, tethered to the downstream T domain, on the thioester carbonyl group of the the amino acid bound to the upstream T domain (Figure 1.19; Challis and Lautru, 2004).

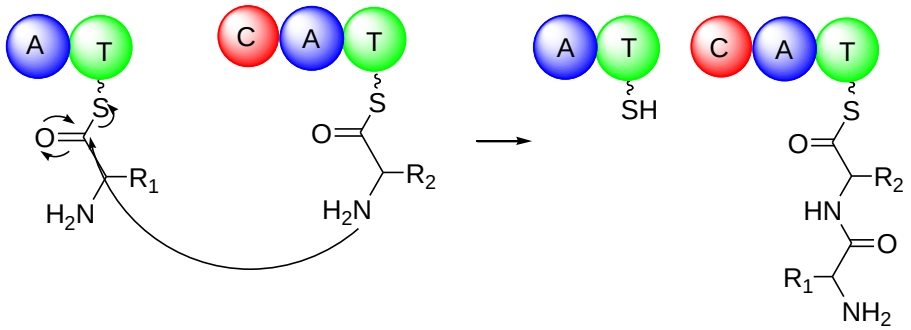


Figure 1.19 The catalytic mechanism of condensation (C) domains.

The final step of polypeptide biosynthesis is release of the peptide from the NRPS, which is usually catalysed by a TE domain located at the C terminal end of the last module of the NRPS. The first structure of a TE domain was solved in 2002 (Brunner et al., 2002). It showed that TE domains contain the catalytic residues Ser80, His207 and Asp107, which catalyse the cleavage of the peptidyl thioester bond resulting in the formation of a peptidyl-O-TE intermediate. This intermediate can be attacked by an external nucleophile, usually a water molecule (hydrolysis) or by an intramolecular nucleophile (cyclization) resulting in the release of the peptide (Figure 1.20; Donadio et al., 2007; Challis and Lautru, 2004).

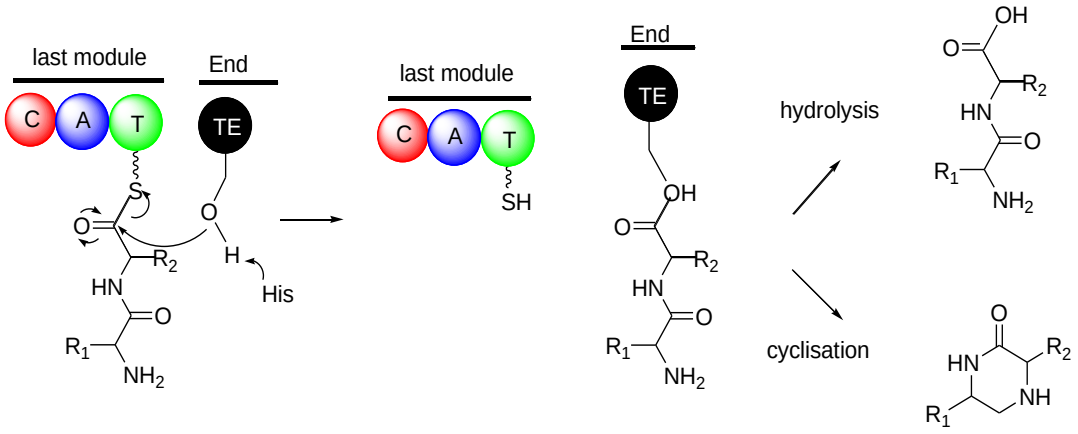


Figure 1.20 The release of peptide from NRPS catalysed by TE domain.

1.5.1.1 NRPS siderophore pathway-Enterobactin Biosynthesis

One of the earliest and most comprehensive studies of an NRPS pathway is the biosynthesis of catecholate siderophore enterobactin (Oves-Costales et al., 2009). The discovery of enterobactin (also named enterochelin) was first reported in 1970 and this compound was found to be a metabolite produced by *Escherichia coli*. The biosynthesis of enterobactin and its ferric complex uptake mechanism have since been elucidated. The total chemical synthesis of enterobactin was reported in 2003

(Raymond *et al*., 2003). The three catecholate moieties in the enterobactin structure allow it to form a very stable ferric complex.

The *entABCDEF* gene cluster of *E. coli* encodes the enzymes responsible for enterobactin biosynthesis. Chorismate, a precursor in aromatic amino acid biosynthesis, is converted into 2, 3-dihydroxybenzoic acid (2, 3-DHB) by EntA, which is the N-terminal domain of EntB and EntC (Rusnak et al., 1990). The three molecules of 2,3-DHB are joined with three molecules of L-serine via three amide and three ester linkages to form enterobactin (Figure 1.21; Liu et al., 1990; Rusnak et al., 1989). EntE is a stand-alone adenylation (A) domain catalysing the acylation of the C-terminal thiolation (T) domain of EntB with 2, 3-DHB (Rusnak et al., 1989; Reichert et al., 1992). The EntF protein contains C, A and T domains and a TE domain is also present at its C-end. The adenylation (A) domain of EntF catalyses the adenylation of the carboxyl group of L-serine and subsequent acylation of the T domain of EntF. The EntF condensation (C) domain catalyses the formation of an amide bond between the amino group in the L-serinyl thioester with the DHB–EntB thioester to form a 2, 3-DHB-L-serinyl-EntF thioesters (Figure 1.21). This is then transacylated with the active Ser residue of the EntF thioesterase (TE) domain to form a 2, 3-DHB-L-serinyl-TE ester (Reichert et al., 1992).

The formation of second 2, 3-DHB-L-serinyl thioester is catalysed by EntF condensation domain and then the TE domain catalyses its condensation with the 2, 3-DHB-L-serinyl-TE ester (Figure 1.22; Shaw-Reid et al., 1999). In this way the dimeric ester is formed. The same process is repeated to form a trimeric ester, which then undergoes macrocyclisation catalysed by the EntF TE domain to release enterobactin (Figure 1.22).

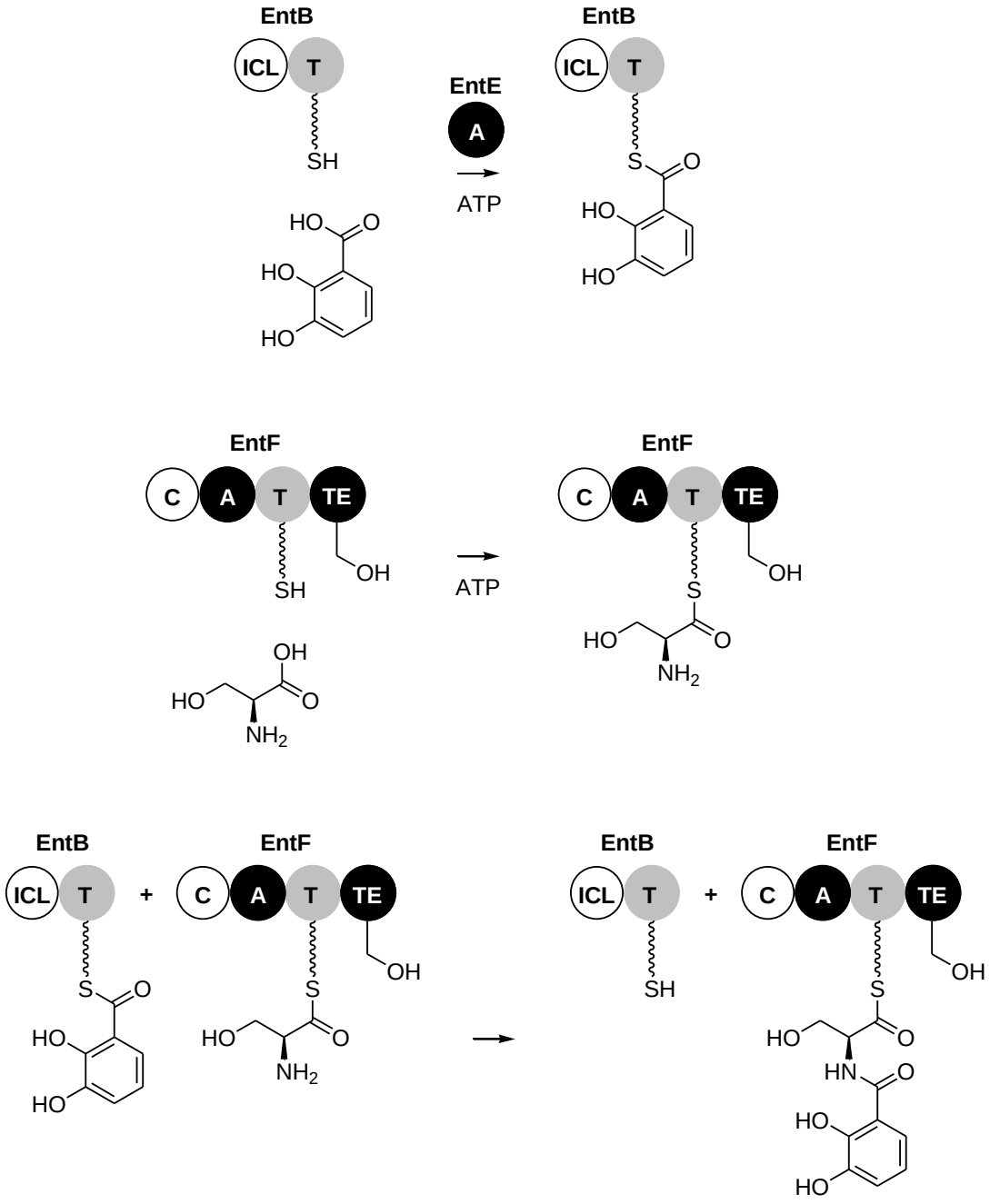


Figure 1.21 Biosynthetic pathway of enterobactin. Reactions catalysed by EntE and EntF enzymes in the assembly of an L-N-(2, 3-dihydroxybenzoyl) serinyl thioester intermediate (Oves-Costales et al., 2009). ICL (EntA)-Isochorismate lyase.

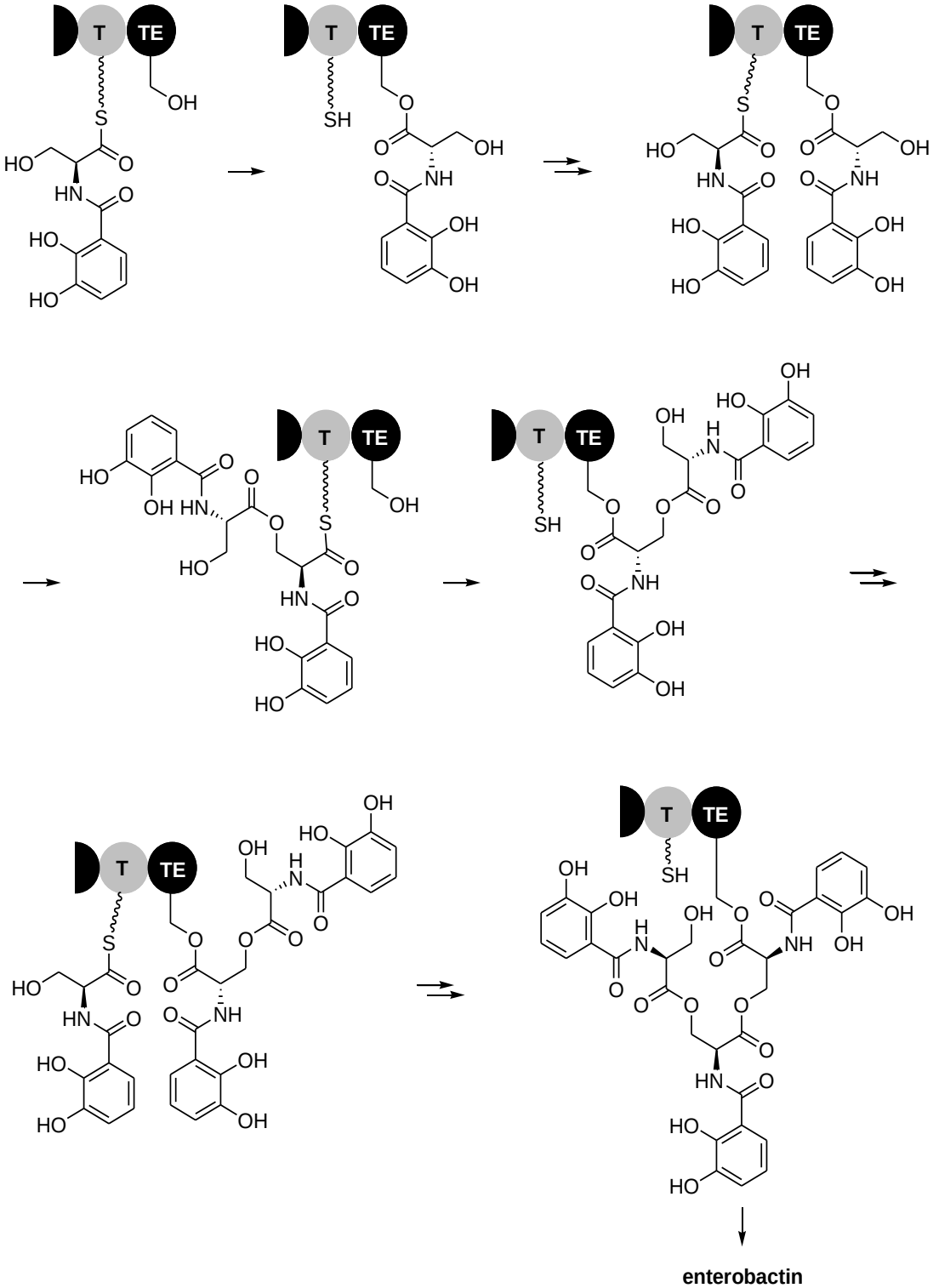


Figure 1.22 Biosynthetic pathway of enterobactin. Oligomerization and macrocyclization reactions catalyzed by the TE domain of EntF from the L-N-(2, 3-dihydroxybenzoyl) serinyl thioester intermediate (Oves-Costales et al., 2009).

1.5.2 Siderophores Assembled via NRPS-Independent Siderophore (NIS) Pathway

Siderophores containing dicarboxylic acid and diamine building blocks are biosynthesised by a structurally distinct family of synthetases, the NRPS-independent siderophore synthetases (NISs). In the last two decades over 40 species of bacteria were found to contain genes encoding for NIS synthetases, which are involved in biosynthesis no less than ten different siderophores (Challis, 2005). The first characterised NIS pathway was the aerobactin biosynthetic pathway (Neilands 1986). Following on from this discovery, several more NIS-derived siderophores were identified including rhizobactin, alcaligin, desferrioxamines, vibrioferrin, staphylobactin, anthrachelin and achromobactin. Genetic studies demonstrated that all these siderophore biosynthetic gene clusters encode at least one enzyme with high sequence similarity to the aerobactin siderophore synthetases. Biosynthesis of siderophores via an NIS pathway involves the activation of carboxylic acid substrates to their corresponding acid adenylates. The NISs can be divided into three subtypes depending on which type of carboxylic acid substrate is recognised and activated (Berti and Thomas, 2009). Type A NISs are specific for citrate, type B for α-ketoglutarate and type C for ester or amid derivatives of carboxylic acid (Oves-Costales et al., 2009). The type A NISs are homologs of the IucA synthetase, type B are homologs of IucD and the type C enzymes are similar to IucC. All of which are present in the aerobactin biosynthetic pathway.

1.5.2.1 Aerobactin Biosynthesis- an archetypal NIS pathway.

Aerobactin is a virulence factor produced by a number of Gram-negative species, including *Escherichia coli*, *Shigella*, *Yersinia* and *Salmonella* (Warner et al., 1981). This siderophore forms a stable hexadentate complex with ferric iron using its carboxylate and hydroxamate groups (Figure 1.23).

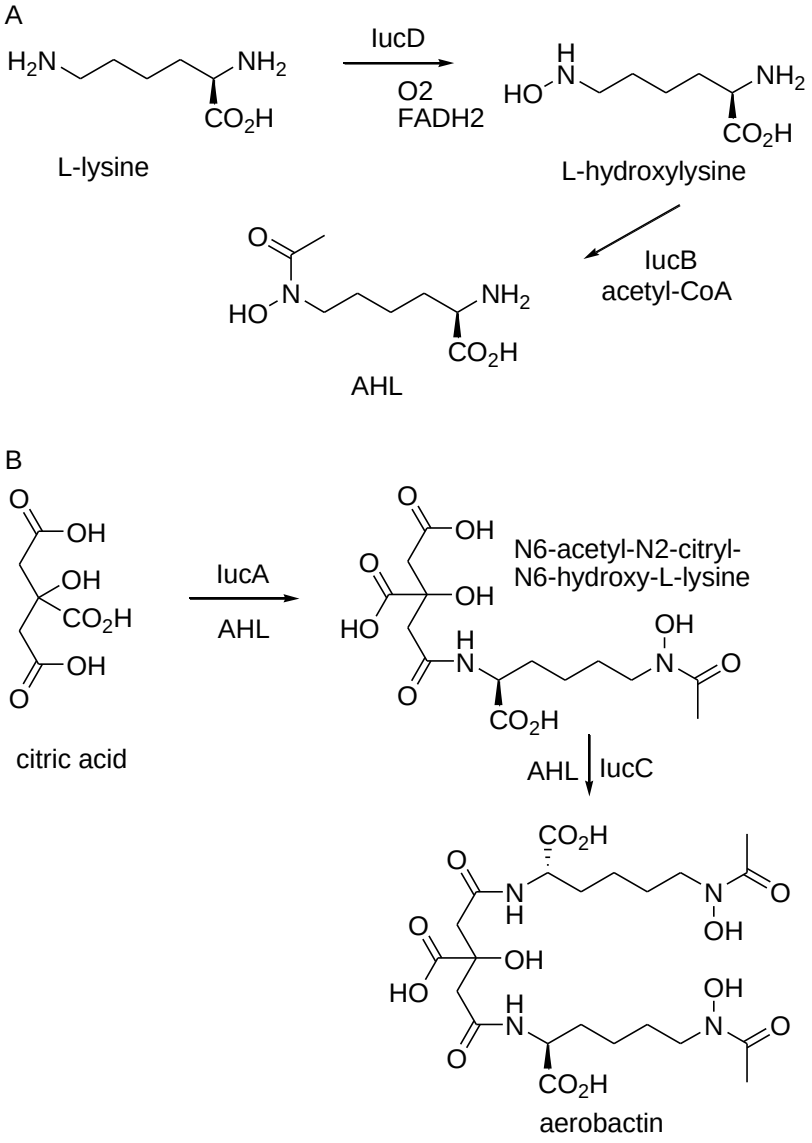


Figure 1.23 Proposed pathway for aerobactin biosynthesis. A-Biosynthesis of N6-acetyl-N6-hydroxylysine (AHL) from L-lysine catalysed by IucB and IucD. B-biosynthesis of aerobactin catalysed by IucA and IucC (Challis, 2005).

The aerobactin biosynthetic gene cluster was first identified on the pColV-K30 plasmid in *E. coli*. It consists of five genes, of which four (*iucABCD*) direct aerobactin biosynthesis (Figure 1.24), while the fifth (*iutA*) encodes an outer-membrane receptor protein involved in the uptake of the ferric-aerobactin complex (Challis, 2005). The *iucD* and *iucB* genes encode a $FADH_2$-dependent monooxygenase and an acetyl-CoA-dependent acyltransferase, respectively. They catalyse the first two steps of aerobactin biosynthesis (Figure 1.23 A). IucD catalyses the conversion of L-lysine to L-N6-hydroxylysine and IucB catalyses the conversion of L-N6-hydroxylysine to L-N6-acetyl-N6-hydroxylysine (AHL; Carbonetti and Williams, 1984; Cox, 1982). The *iucA* is proposed to encode a synthetase catalysing the condensation of the α-amino group in L-N6-acetyl-N6-hydroxylysine with one of the prochiral carboxyl groups of citric acid to form N6-acetyl-N2-citryl-N6-hydroxy-L-lysine (Figure 1.24). The *iucC* is proposed to encode a homologue of IucA synthetase, which catalyses the condensation of the product of the IucA-catalysed reaction with a second molecule of L-N6-acetyl-N6-hydroxylysine to form aerobactin (Challis, 2005; Figure 1.23).

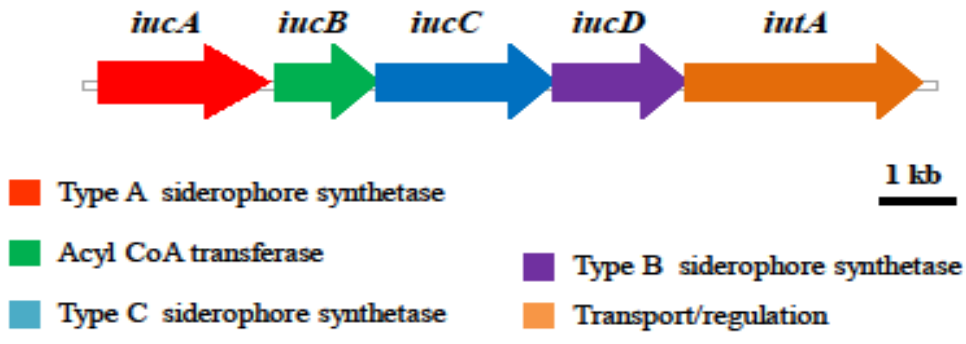


Figure 1.24 Aerobactin biosynthetic gene cluster.

1.6 Investigations of Cryptic Siderophore Biosynthetic Pathways encoded for in the *S. scabies* 87.22 genome.

Not much is known about the significance of siderophores in bacterial plant pathogenicity. Investigations into siderophore mediated iron acquisition and its role in plant pathogenicity has only been explored in Gram-negative bacteria known to be plant pathogens. For example, the siderophore pyoverdine plays a key role in the infection of tobacco by *Pseudomonas syringae* pv. tabaci 6605 (Sriyosachati and Cox, 1986). The presence of iron aids the transcription of virulence genes during speck disease in tomato by *Pseudomonas syringae* pv. tomato DC3000 (Kim et al., 2009). However, to date, siderophore-mediated iron uptake and its role in plant pathogenicity has not been explored as extensively in Gram-positive bacteria. One Gram-positive strain, where studies into this have been undertaken, is the plant pathogen *S. scabies* 87.22.

1.6.1 Pyochelin

Analysis of the *S. scabies* 87.22 genome (http://strepdb.streptomyces.org.uk/) uncovered a putative gene cluster for the biosynthesis of pyochelin, a previously characterised siderophore. Pyochelin (Figure 1.25) produced by the mammalian pathogen *Pseudomonas aeruginosa* (Cox et al., 1981) is known to be a virulence factor in these bacteria, as it plays an important role in the infection in mice (Cox, 1982). This is because pyochelin is able to take up iron from transferrin and lactoferrin present in the blood serum of mammals and use it for bacterial growth (Akenbauer et al., 1995).

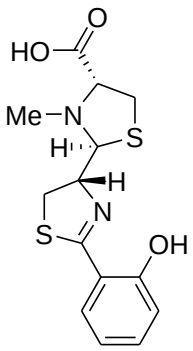


Figure 1.25 Structure of pyochelin (Cox, 1982).

1.6.1.1 Pyochelin Biosynthesis in *P. aeruginosa*

Pyochelin is a siderophore that contains a thiazoline/thiazolidine moiety and its biosynthesis in *Pseudomonas aeruginosa*, has been fully characterised (Cox, 1982; Takase et al., 2000). The pyochelin biosynthetic gene cluster in *P. aeruginosa* (Figure 1.26) encodes seven proteins PchABCDEFG, which are required for pyochelin biosynthesis (Figure 1.27; Cox et al., 1981; Reinmann et al., 2001).

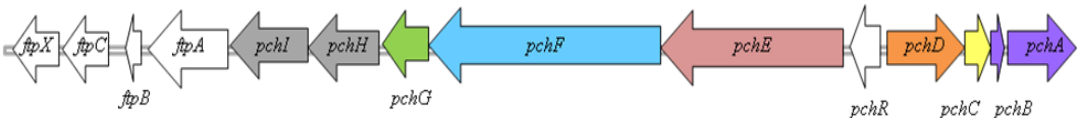


Figure 1.26 Organization of the pyochelin biosynthetic gene cluster in *P. aeruginosa* PAO1.

PchA and PchB catalyse the biosynthesis of salicylate from chorismate (Serino et al., 1995). A stand-alone A domain, PchD, then adenylates salicylate and subsequently loads it onto the thiolation domain of PchE NRPS multienzyme (Figure 1.27). The A domain of PchE catalyses the acylation of cysteine onto the C-terminal T domain, and the Cy (heterocyclization) domain catalyses condensation of the cysteinyl thioester with the salicylic thioester to form a peptide bond. An MT-like domain

embedded within the A domain of PchE is then proposed to catalyse the epimerization of the α-carbon of the resulting thioester (Patel *et al.*, 2003). The Cy domain, then, catalyses the cyclodehydration to yield a 2-hydroxyphenyl-thiazolinyl thioester intermediate.

The A domain of PchF catalyses the *cis* loading of cysteine onto its T domain, and the Cy domain of PchF catalyses amide bond formation between the cysteinyl thioester and the 2-hydroxyphenyl-thiazolinyl thioester on PchE. Subsequent cyclodehydration results in a formation of a bis-thiazolinyl thioester intermediate. PchG catalyses NADPH-dependent reduction of one of the thiazoline rings to the corresponding thiazolidine and an MT domain in PchF catalyzes the S-adenosylmethionine (SAM)-dependent N-methylation of the thiazolidine. The TE domain of PchF catalyses hydrolysis of the thioester bond in the resulting intermediate yielding pyochelin (Reinmann et al., 2001; Serino et al., 1995; Serino et al., 2003).

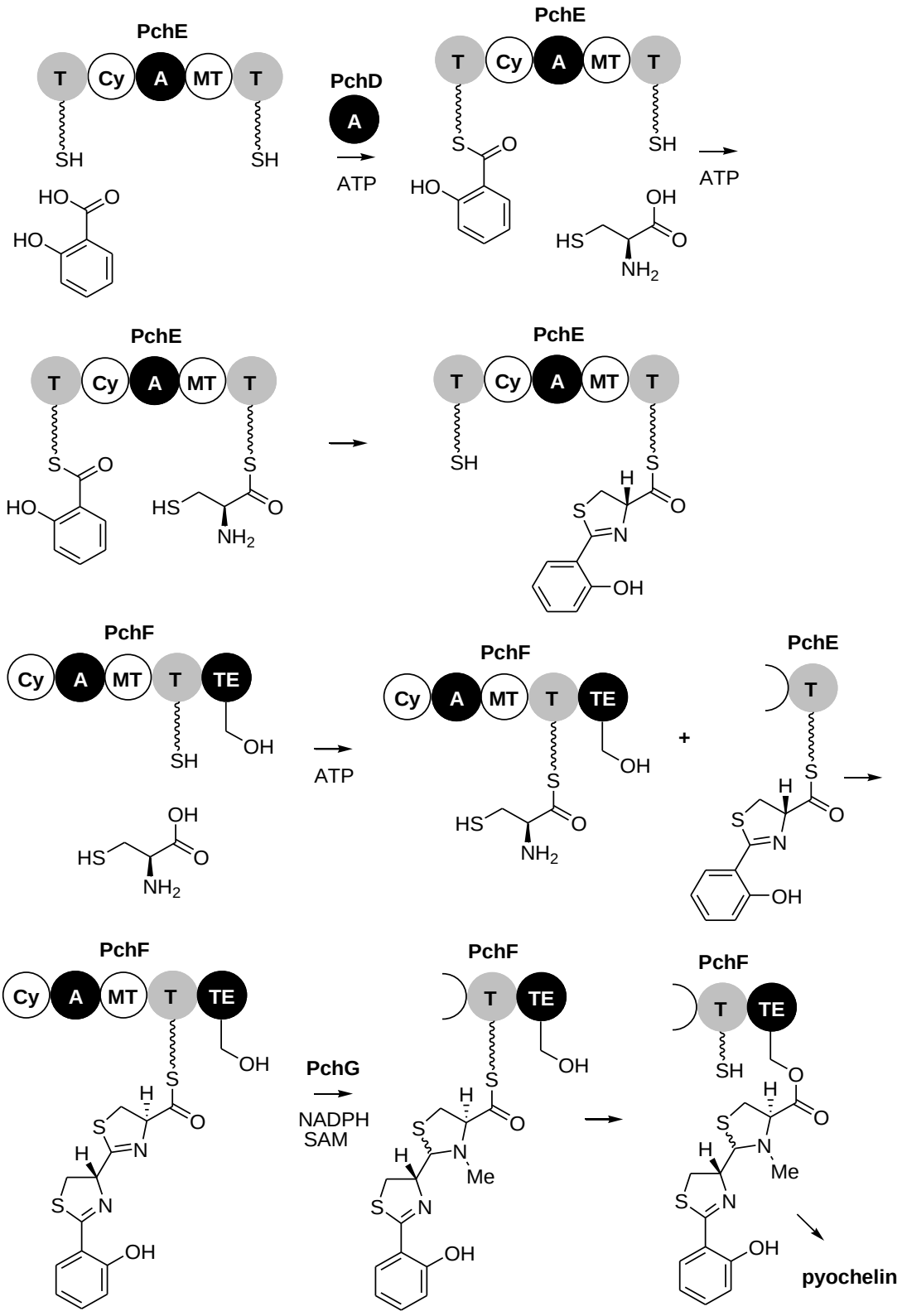


Figure 1.27 Biosynthesis of pyochelin. NRPS domains: T-thiolation, Cy-heterocyclization, A-adenylation, MT-methyltransferase, TE-thioesterase (Reinmann et al., 2001; Serino et al., 1995; Serino et al., 2003).

1.6.1.2 Sequence Analysis of the *S. scabies* Pyochelin Biosynthetic Gene Cluster.

The pyochelin biosynthetic gene cluster (PBGC) in *S. scabies* contains 12 genes (Figure 1.28; Cox, 1982). Six of these genes encode proteins (Scab1381, Scab1411, Scab1421, Scab1461, Scab1471 and Scab1481) that are proposed to be analogous in function to the pyochelin biosynthetic proteins PchABCDEFG from *P. aeruginosa* (Patel and Walsh, 2001; Quadri et al., 1999; Serino et al., 1997).

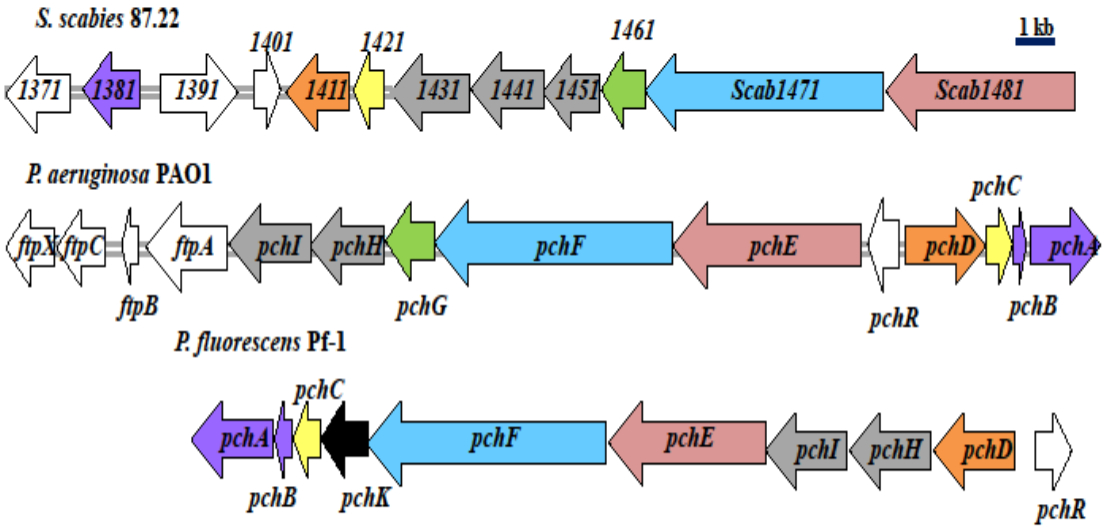


Figure 1.28 Organisation of pyochelin biosynthetic gene clusters in *S. scabies* 87.22, *P. aeruginosa* PAO-1and *P. fluorescens* Pf-5 (Seipke et al., 2011; Youard et al., 2007).

Scab1381 encodes for a putative salicylate synthase and it contains an anthranilate synthase domain. Scab1411 is predicted to be AMP-ligase, which contains putative active site residues that are highly similar to the active site residues of MbtA, an AMP-ligase from *Mycobacterium tuberculosis*. MbtA catalyses the adenylation of salicylate in mycobactin biosynthesis (Challis et al., 2000; Quadri et al., 1998; Rausch et al., 2005). Two genes *scab1481* and *scab1471*, encode proteins with putative functions analogous to PchF and PchE, respectively, both of which are

NRPS enzymes in *P. aeruginosa* (Figure 1.28). Scab1481, which is analogous to PchE, is proposed to catalyse the formation of a 2-hydroxyphenylthiazolinyl thioester from the salicyl adenylate and Scab1471 adds a second molecule of the cysteine to the thioester to form 2-hydroxyphenylbis-thiazolinyl thioester attached to the T domain of Scab1471 in a manner analogous to PchF (Figure 1.27). *Scab1461* is predicted to encode a reductase enzyme homologous to PchG. It is proposed to catalyse the reduction of one of the thiazoline rings in the thioester. N-methylation of the reduced thiazoline ring and the release of the pyochelin by the N-terminal thioesterase domain is catalysed by Scab1471 (Patel and Walsh, 2001). Scab1421 protein is predicted to be a homolog of PchC, a proofreading thioesterase, which catalyses removal of wrongly charged molecules from the PchE and PchF thiolation domains (Reinmann et al., 2004). The analysis of the *S. scabies* genes indicates that they encode proteins, which have identical properties to the pyochelin biosynthetic proteins PchABCDEFG from *P. aeruginosa* and indicates that *S. scabies* may be a producer of pyochelin.

Other genes from the putative pyochelin gene cluster, such as *scab1401* encodes a pathway specific regulator homologous to PchR, a pyochelin-responsive repressor protein in *P. aeruginosa* and Scab1401 is a TetR family transcriptional regulator. Analysis of the transcription of the PBGC revealed that *scab1371* is a positive transcriptional regulator of the cluster and it is co-transcribed with the *scab1381*, predicted to encode salicylate synthase. Putative transport proteins are encoded for by *scab1431*, *scab1441* and *scab1451*.

There are also some differences between the biosynthetic proteins in *P. aeruginosa* and *S. scabies*. Formation of salicylate from chorismate is catalysed in *P. aeruginosa* by two proteins PchA and PchB (Serino et al., 1995) while in *S. scabies* 87.22 a single protein, Scab1381, is predicted to catalyse the same reaction. Another difference is that the *S. scabies* pyochelin gene cluster does not contain an ortholog of the *P. aeruginosa fptX* gene encoding the inner membrane pyochelin permease. The gene is not encoded anywhere else in the *S. scabies* genome and therefore, it is speculated that one of the putative transport proteins Scab1431, Scab1441 or Scab1451 within the cluster may be used for the import of ferric-pyochelin in *S. scabies* (Seipke et al., 2011).

1.6.1.3 Biosynthesis of Enantio-Pyochelin in *Pseudomonas fluorescens*

Pyochelin is biosynthesised from salicylate and two molecules of cysteine. During thiazoline ring formation in *P. aeruginosa*, the α-carbon of the first cysteine residue undergoes epimerisation, which results in the conversion of cysteine to its D-isoform, whereas the second cysteine remains as L-isoform (Figure 1.27). This results in the formation of two pyochelin diastereoisomers with the stereochemistry 4'R, 2''R, 4''R for pyochelin I and 4'R, 2''S, 4''R for pyochelin II (Youard et al., 2007). Other closely related strains *P. fluorescens* Pf-5 and CHA0 were found to produce an enantiomer of pyochelin named enantiopyochelin (Figure 1.29). Enantiopyochelin was found to facilitate growth under iron-deficient conditions in *P. fluorescens* and to induce the expression of its biosynthetic genes, however, it did not promote growth in *P. aeruginosa.* On the other hand, pyochelin from *P. aeruginosa* did not promote growth, nor induce the expression in *P. fluorescens*. Comparison of

the pyochelin biosynthetic gene cluster between *P. aeruginosa* and *P. fluorescens* (Figure 1.28) highlighted three main differences. Firstly, in *P. aeruginosa* the biosynthetic genes are organised in two operons, while in *P. fluorescens* the genes are in one operon. Secondly, PchE in *P. fluorescens* does not contain the epimerisation domain and finally the predicted function of PchK in *P. fluorescens* is not homologous to PchG in *P. aeruginosa*. PchG is a thiazoline reductase and PchK is predicted to have reductase and/or epimerase function. The amino acid sequences of other of *P. aeruginosa* and *P. fluorescens pch* genes are between 35 and 60% identical (Youard et al., 2007).

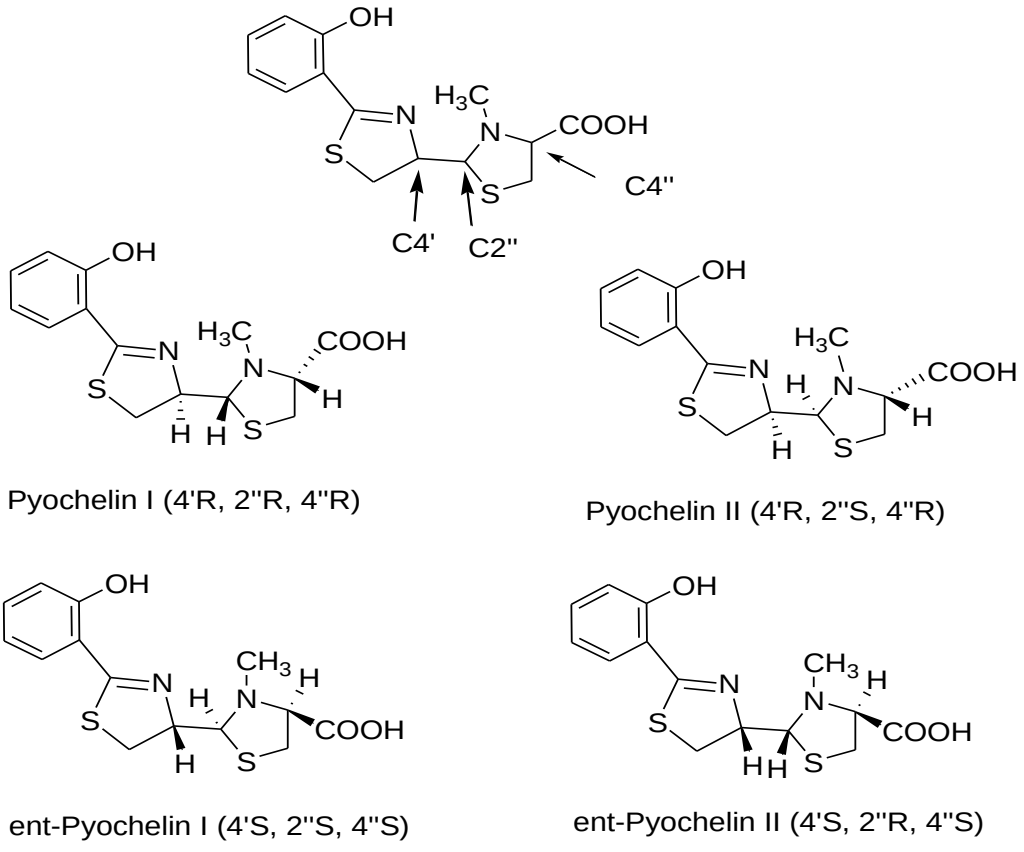
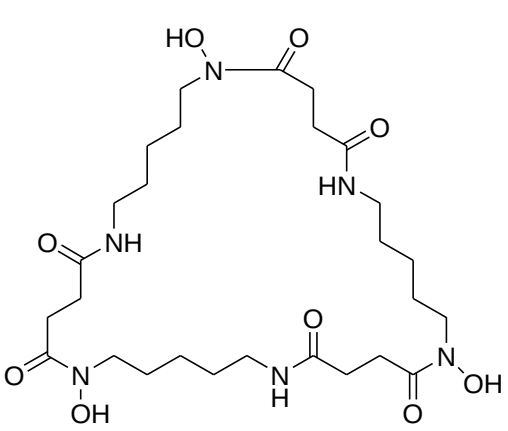


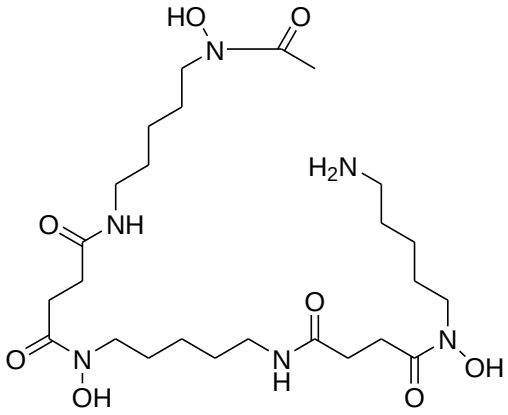
Figure 1.29 Stereoisomers of pyochelin (*P. aeruginosa*) and ent-pyochelin (*P. fluorescens*).

1.6.2 Desferrioxamines.

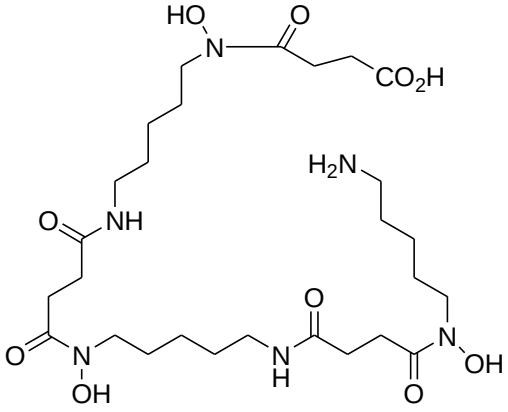
Desferrioxamines are a group of siderophores containing three hydroxamic acid ligands, which form a hexadentate complex with ferric iron (Miller, 1989). Desferrioxamines include the linear trimer desferrioxamine G1, the hetero-oligomer desferrioxamine B and the 33-membered macrocyclic desferrioxamine E (Figure 1.30). They are biosynthesised by several *Streptomyces* species (Bickel et al., 1960). *S. coelicolor* has been shown to produce desferrioxamine G_1 and E (Imbert et al., 1995) and *S. viridosporus* has been shown to produce desferrioxamine B and E (Imbert et al., 1995). *S. pilosus* has been found to take up ferrioxamines B and E, *S. lividans* B and G_1 and *S. viridosporus* B, E and G_1 (Muller and Raymond, 1984, Imbert et al., 1995). Desferrioxamine B is used clinically for the treatment of iron poisoning in human body.



Desferrioxamine E



Desferrioxamine B



Desferrioxamine G1

Figure 1.30 Structure of desferrioxamines E, B and G1.

Analysis of the *S. coelicolor* M145 genome sequence identified a cluster of six genes (*desEFABCD*) that direct the biosynthesis of desferrioxamines E and B (Figure 1.31). The Des cluster is also conserved in *S. avermitilis* and *S. ambofaciens* (Barona-Gomez et al., 2004).

Analysis of the *S. scabies* 87.22 genome sequence resulted in the identification of a cryptic biosynthetic gene cluster, which includes the protein coding sequences Scab57921–Scab57981 (Figure 1.31). These proteins are homologs of proteins

encoded for by genes in the desferrioxamine biosynthetic gene cluster of *S. coelicolor* A(3)2 (Table 1.3). Preliminary LC-MS analysis of *S. scabies* culture supernatants indicated that *S. scabies* also produces desferrioxamines (Song and Challis, unpublished data).

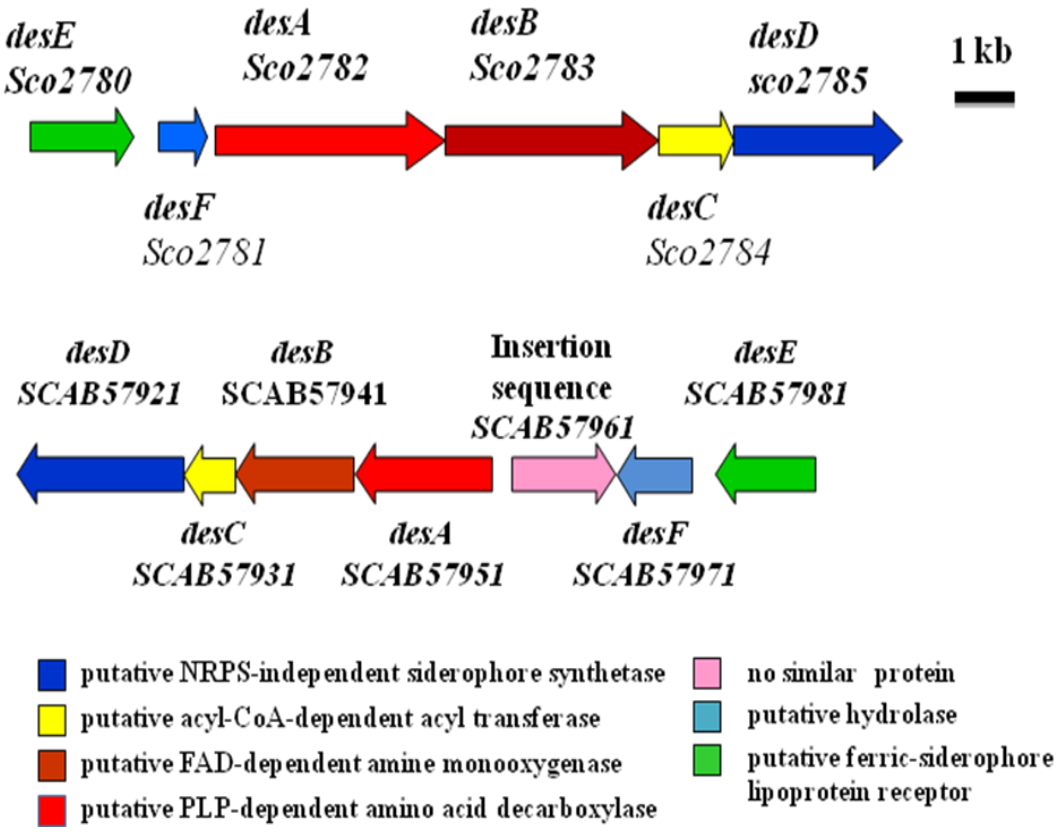


Figure 1.31 Organistation of the gene cluster directing desferrioxamine biosynthesis and ferrioxamine uptake in *S. coelicolor* M145 and in *S. scabies* 87.22.

Table 1.3 Putative proteins function encoded by the desferrioxamine biosynthetic gene cluster in *S. scabies* 87.22

***S. scabies* gene**	**desferrioxamine biosynthesis protein (Des)**	**Sequence similarity with *S. coelicolor* A3(2) *des* cluster (%)**
scab57921	DesD	hypothetical protein SCO2785 (81%)
scab57931	DesC	aceytltranferase SCO2784 (79%)
scab57941	DesB	monooxygenase SCO2783 (80%)
scab57951	DesA	pyridoxal-dependent decarboxylase SCO2782 (86%)
scab57961	IS630 family insertion sequence	no similar protein
scab57971	DesE	hypothetical protein SCO2781 (75%)

The first steps of desferrioxamine biosynthesis were investigated in *Streptomyces pilosus*, and involved elucidation of reactions catalysed by the enzymes DesA and DesB (Schupp at al., 1987, 1988), which have high sequence identity to PLP-dependent amino acid decarboxylases and FAD-dependent amine monooxygenases, respectively. DesA and DesB are involved in the first two steps of desferrioxamine biosynthesis to form N-hydroxycadaverine from L-lysine (Figure 1.32 A).

Investigations into the roles of DesC and DesD were carried out in the desferroxamine E biosynthetic pathway of *S. coelicolor*. The *desC* gene encodes an enzyme similar to acyl-CoA-dependent acyl transferases (Kadi et al., 2007) and is proposed to catalyse the N-acylation of *N*-hydroxycadaverine (N. Kadi and G. L. Challis, unpublished data). DesC is predicted to have relaxed substrate specificity in

those bacteria, which also produce desferrioxamine B, catalysing acylation of N-hydroxycadaverine with acetyl-CoA as well as succinyl-CoA, giving N-hydroxy-N-acetylcadaverine (HAC) and N-hydroxy-N-succinylcadaverine (HSC), respectively (Figure 1.32 A; Barona-Gomez et al., 2004).

The gene *desD* encodes a protein that has a high percentage sequence similarity to IucA and IucC from the aerobactin biosynthetic pathway, the first NIS enzymes to be characterised. Purified recombinant DesD was shown to catalyse the ATP-dependent conversion of chemically-synthesized HSC to desferrioxamines G1 and E. DesD catalyses the ATP-dependent trimerisation and macrocyclisation of HSC to yield desferrioxamine E (Figure 1.32 B) and the ATP-dependent condensation of two units of HSC with one unit of HAC to yield desferrioxamine B (Figure 1.32C; Barona-Gomez et al., 2006).

DesE is similar to ferric-siderophore lipoprotein receptors and has been shown to bind ferrioxamines with high affinity and selectivity (Patel et al., 2010). DesF is similar to ViuB, which is proposed to be a hydrolase involved in the release of iron from ferric-vibriobactin. However, more recent sequence analyses suggest it is a ferric-siderophore reductase (Barona-Gomez et al., 2006).

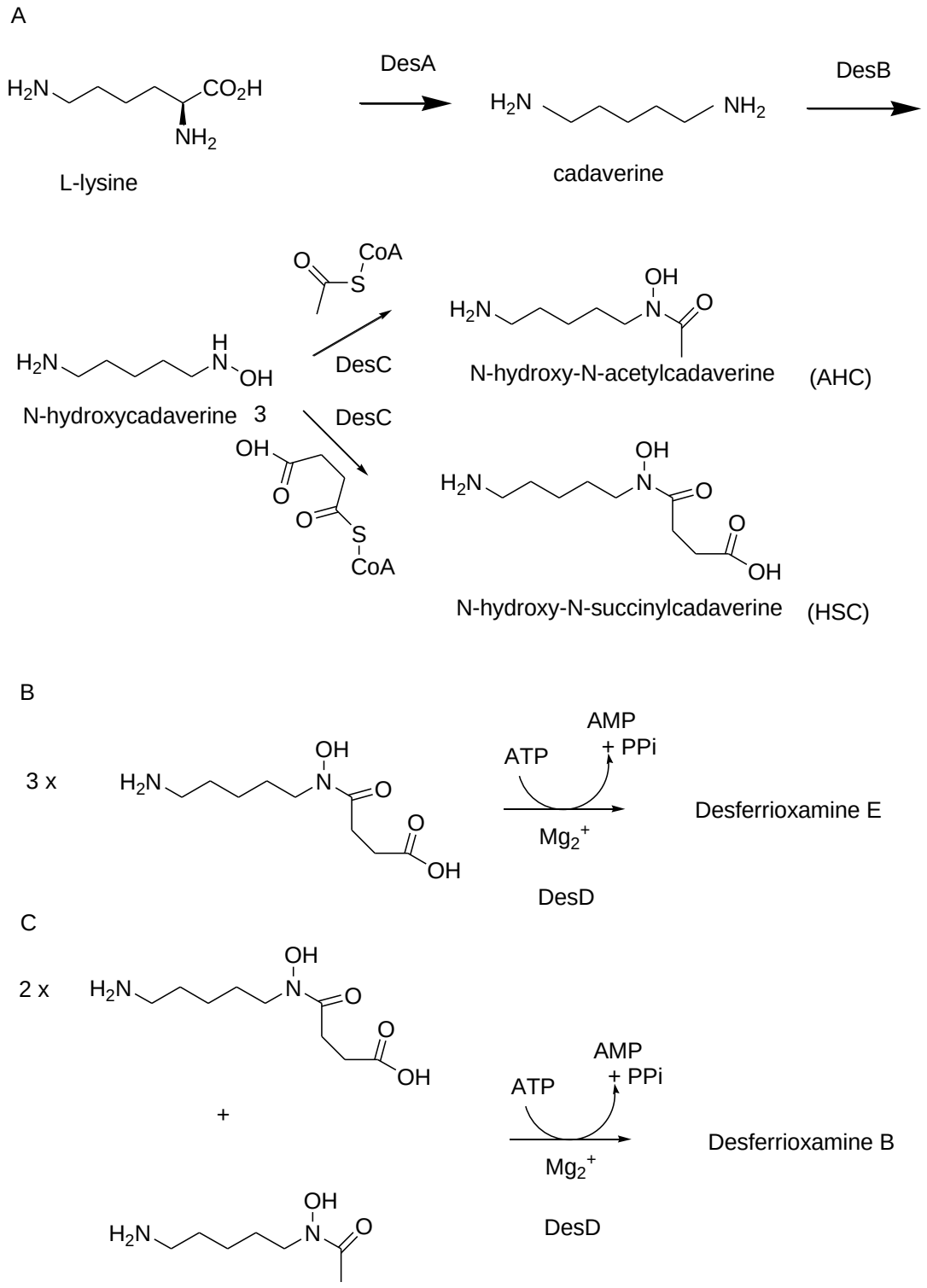


Figure 1.32 Proposed pathways for desferrioxamine E and B biosynthesis. Proposed pathway for the biosynthesis of AHC and HSC. B-Biosynthesis of desferrioxamine E by DesD. C- Biosynthesis of desferrioxamine B by DesD (Kadi et al., 2007).

1.6.3 Novel Putative Hydroxamate Siderophore

A novel tris-hydroxamate iron-chelating metabolite was originally discovered in *Streptomyces antibioticus* NBRC 13838 by Dr Shinya Kodani and coworkers (Shizuoka University, Japan), collaborators of the Challis group. Analysis of the *S. scabies* 87.22 genome sequence unveiled a cryptic NRPS biosynthetic gene cluster (Figure 1.33), which includes the protein coding sequences from Scab85431 to Scab85521, suggesting that it may encode an NRPS that directs the production of a hydroxamate-containing metabolite. All of these proteins, present within the NRPS biosynthetic gene cluster, are homologous to proteins encoded for by genes within the coelichelin biosynthetic gene cluster from *S. coelicolor* M145 (Table 1.4) (http://www.ncbi.nlm.nih.gov/nuccore/260644157).

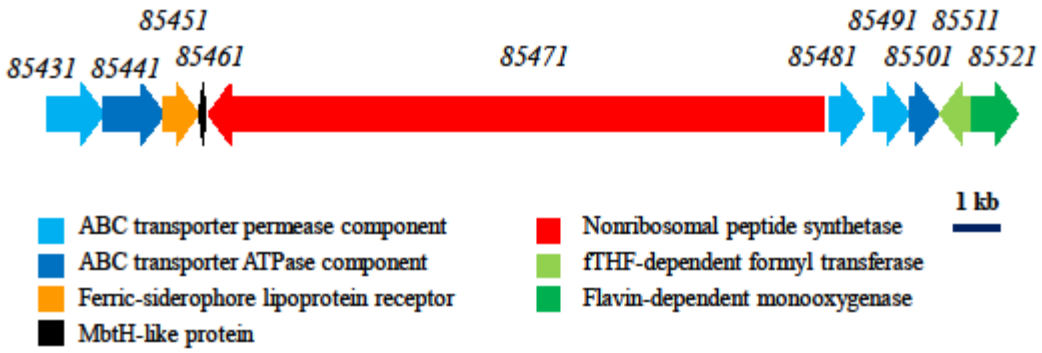


Figure 1.33 Organization of the scabichelin biosynthetic gene cluster in *S. scabies* 87.22.

Table 1.4 Putative proteins function encoded by the putative novel hydroxamate siderophore biosynthetic gene cluster in *S. scabies* 87.22.

Protein	**Putative function**	**Homolog; origin**	**Similarity/ identity (%)**
Scab85431	ABC transporter permease/ATPase component	CchI; *S. coelicolor*	53/38
Scab85441	ABC transporter permease/ATPase component	CchG; *S. coelicolor*	58/46
Scab85451	Ferric-siderophore lipoprotein receptor	CchF; *S. coelicolor*	53/38
Scab85461	MbtH-like protein	CchK; *S. coelicolor*	72/57
Scab85471	Nonribosomal peptide synthetase	CchH; *S. coelicolor*	60/48
Scab85481	ABC transporter permease component	CchC; *S. coelicolor*	65/47
Scab85491	ABC transporter permease component	CchD; *S.coelicolor*	60/45
Scab85501	ABC transporter ATPase component	CchE; *S. coelicolor*	74/60
Scab85511	Formyl-tetrahydrofolate-dependent formyl transferase	CchA; *S. coelicolor*	85/76
Scab85521	Flavin-dependent monooxygenase	CchB; *S. coelicolor*	73/60

The genes *Scab85511* and *scab85521* encode proteins homologous to CchA and CchB from *S. coelicolor*, respectively. CchA and CchB are involved in the biosynthesis of the non-proteogenic amino acid residues hydroxyornithine (hOrn) and formylhydroxyornithine (fhOrn), which are take part in the biosynthesis of

coelichelin. An NRPS multienzyme is encoded for by *scab85471* and is proposed to be involved in the biosynthesis of a putative novel hydroxamate siderophore (Figure 1.34; Lautru and Challis, 2004).

Bioinformatic analysis of the functional domains of the NRPS encoded for by *scab85471* revealed it contains 17 enzymatic domains assembled into 5 modules (Figure 1.34). Each module contains an adenylation (A), a thiolation (T) and a condensation domain (C) (http://www.nii.res.in/searchall.html). The final termination module of *scab85471* also contains an epimerisation/condensation domain (Figure 1.34) suggesting that the stereochemistry of the peptide is L-L-L-L-D. It is hypothesised that product of the *scab85471* gene would be a pentapeptide, as it contains 5 modules. The substrate specificity of the adenylation domains in each module was predicted using the NRPS predictor online tool (Table 1.5; Raush et al., 2005). It was predicted that modules 1, 2 and 3 of the NRPS probably have specificity for a L-N5-acyl-N5-hydroxyornithine, a L-Ser and a L-N5-hydroxyornithine, respectively. The substrate specificity for the modules 4 and 5 could not be predicted, indicating that unusual substrates may be used by these modules. Modules 1 and 3 of the NRPS contain N-methyltransferase (MT) domains suggesting that residues 1 and 3 could be N-methylated.

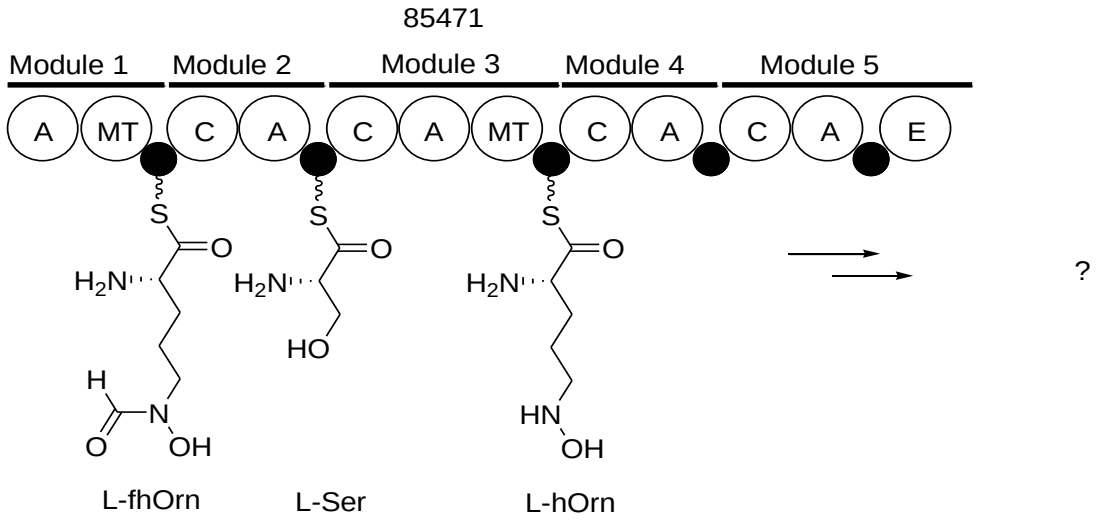


Figure 1.34 Organisation of modules and domains of the NRPS encoded by *scab85471*.

Table 1.5 Predicted substrate specificity of adenylation domains in the Scab85471 protein.

Module number	**Substrate specificity**	**Residue position**
1	hforn	DVWILGAT
2	ser	DVWHLSLV
3	horn	DMENLGLI
4	?	DGEDLGLS
5	?	DAQXXXLV

1.7 Products of Other Cryptic Pathways

A family of polyketide natural products, which all contain a five-membered tetronate ring are known to have apoptotic and antibacterial properties (Hamaguchi et al., 1995). All these natural products were found to be biosynthesised by both Gram-negative and Gram-positive bacteria, including various *Streptomyces* species.

Examples of tetronate containing natural products are RK-682 (Figure 1.35) from *Streptomyces sp.* 88-68 and agglomerin antibiotics (Figure 1.35) from the Gram-negative *Pantoea agglomerans* (Sun et al., 2010).

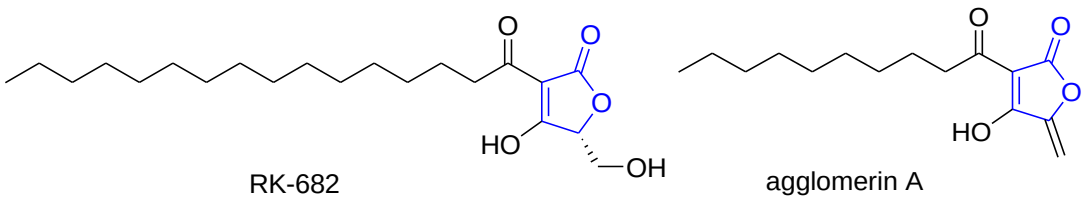


Figure 1.35 Structures of tetronate polyketides.

The common structural feature of tetronate natural products is the tetronic acid ring (Coloured blue in Figure 1.35). The mechanism of tetronic acid formation was elucidated during investigations into the biosynthesis of RK-682. This is proposed to involve an FabH-like protein, which is an enzyme belonging to a family of fatty acid synthases (Sun et al., 2010).

Another common feature of these natural products is a polyketide backbone, which is assembled by enzymes belonging to the family of polyketide synthases (PKSs). Polyketides comprise a major group of structurally and functionally diverse natural products that include a broad range of clinically important compounds, such as antibiotics, immunosuppressants, antitumour, antifungal and antiparasitic agents (McMurry and Begley, 2005). Polyketides are biosynthesized by the condensation of extender units malonyl-, methylmalonyl- or ethylmalonyl-CoA with an acyl-CoA starter unit and are classified according to their structure into three types. Type I PKSs contain multiple catalytic domains for each reaction in the polyketide chain

assembly and are classified into modular and iterative systems. In type I modular PKSs each module is used only once to catalyse the elongation of the polyketide chain by a malonate extender unit, whereas in iterative systems the PKS modules are used for more than one elongation step. Type II PKSs are a group of monofunctional proteins where each protein catalyses a different reaction in the chain assembly. Type III PKSs are homodimers of identical keto synthase (KS) domains that do not use ACP domains in the polyketide chain assembly. They posses a single active site to catalyse different types of reactions, such as decarboxylations, condensations and cyclisations (Austin and Noel, 2002).

Type I PKSs are involved in tetronate polyketide chain assembly. In a type I PKS the loading module contains an acyl transferase (AT) and an acyl carrier protein (ACP) A chain extension module consists of a minimum of three domains: an acyl transferase (AT), an acyl carrier protein (ACP) and a β-ketoacyl synthase (KS). In the first stage, the AT domain of the loading module loads the acyl-CoA starter unit and then transfers it to the adjacent ACP domain (Figure 1.36 A; McMurry and Begley, 2005). During chain extension the malonate extender unit is loaded onto the AT domain of the module 1 and then transferred onto the ACP domain within the same module (Figure 1.36 B; McMurry and Begley, 2005). The KS domain of module 1 then catalyses the decarboxylative Claisen condensation between the ACP bound acyl thioester of the module 1 with the acyl-ACP of the upstream loading module to extend the chain forming a β-ketothioester (Figure 1.37; McMurry and Begley, 2005). Additional domains, such as a ketoreductase (KR), a dehydratase (DH) and an enoyl reductase (ER) may be present in some extension modules. The KR domain catalyses the NADPH-dependent reduction of the β-carbon to form a β-

hydroxy thioester (Figure 1.38; McMurry and Begley, 2005). The DH domain catalyses the conversion of the β-hydroxy thioester to an α, β-unsaturated thioester by removal of water (Figure 1.39; McMurry and Begley, 2005). The ER domain catalyses the NADPH-dependent reduction of the α, β-unsaturated double bond to form a saturated bond (Figure 1.40; McMurry and Begley, 2005).

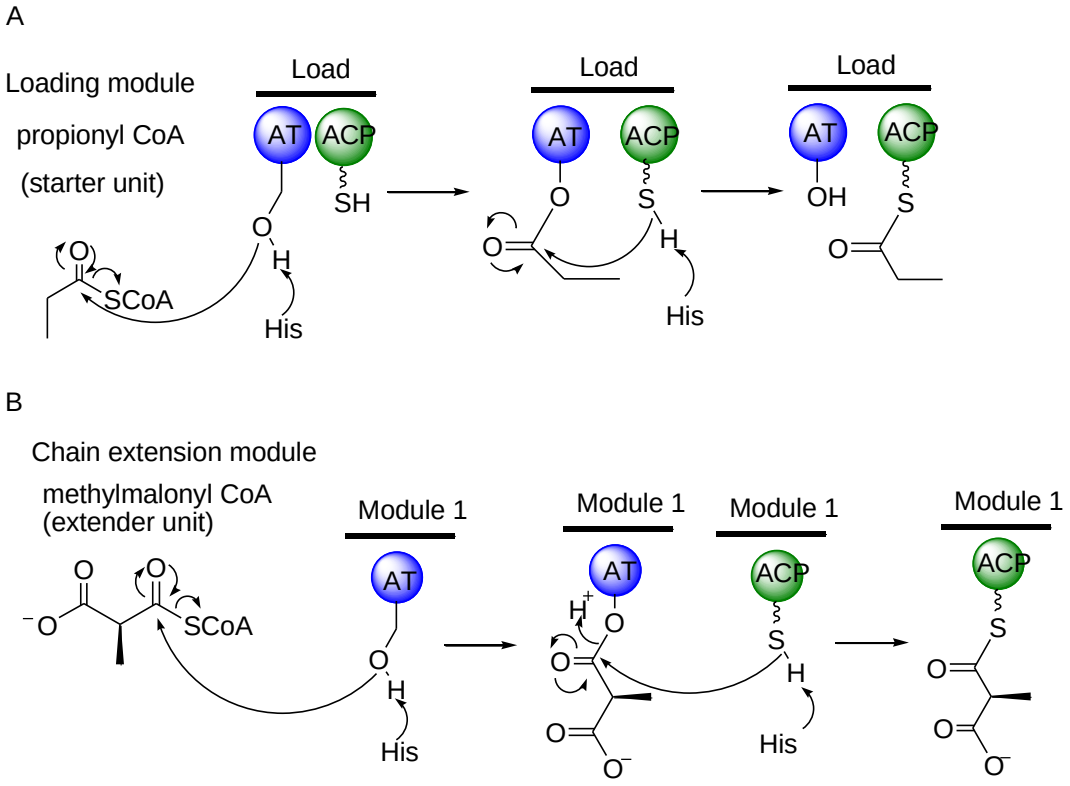


Figure 1.36 Reaction catalysed by AT domains. A-loading, B-chain extension

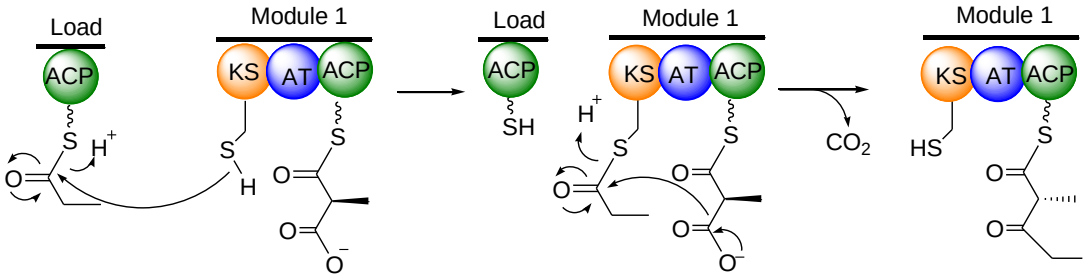


Figure 1.37 Reaction catalysed by KS domain.

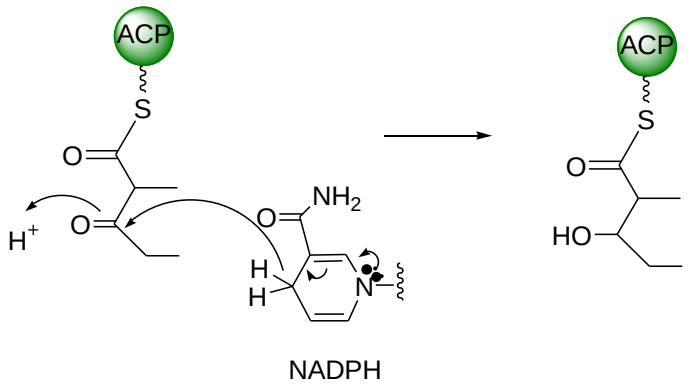


Figure 1.38 Reaction catalysed by KR domain.

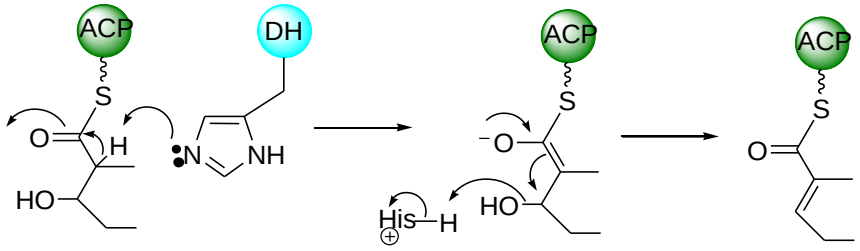


Figure 1.39 Reaction catalysed by DH domain.

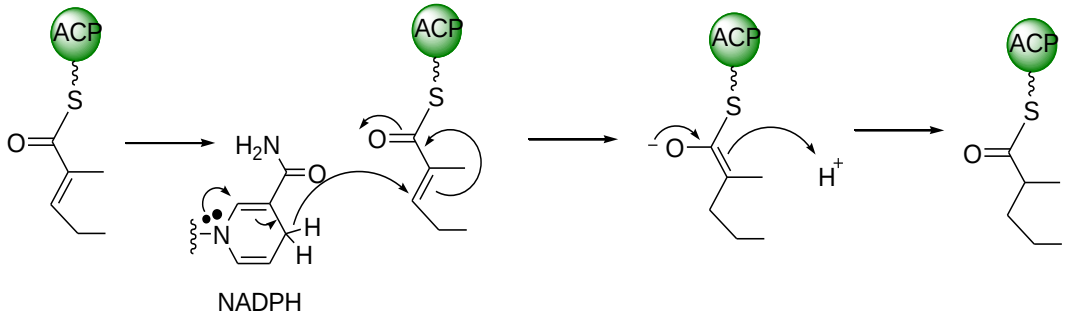


Figure 1.40 Reaction catalysed by ER domain.

1.7.1 RK-682

RK-682 (3-hexadecanoyl-5-hydroxymethyl-tetronic acid; Figure 1.35) is a model tetronate containing polyketide, originally isolated from *Streptomyces* sp. 88-682 (Hamaguchi et al., 1995). It is an inhibitor of protein phosphatases and of HIV-1 proteinase. The RK-682 biosynthetic gene cluster was heterologously expressed in *Saccharopolyspora erythraea* JCB2, which does not produce erythromycin, thus allowing the production of RK-682, to be easily detected (Sun et al., 2010).

Nine genes, RkA-RkI, were identified in the RK-682 biosynthetic gene cluster in *Streptomyces* sp. 88–682 (Figure 1.41). The genes from this cluster code for the following proteins: *rkI*-TetR-family transcriptional regulator; *rkH*- antibiotic efflux protein; *rkA*- palmitoyl-CoA synthetase; *rkB* and *rkF*- acyl carrier proteins; *rkC*, modular PKS containing KS, AT and ACP domains; *rkD*-FabH-like protein; *rkE*-FkbH-like glyceryl-*S*-ACP synthase; *rkG* -type II thioesterase (TEII). It was determined that expression of all nine genes from the RK-682 cluster from the original host is required for similar production levels in *Saccharopolyspora erythraea*. The putative Rk proteins were later expressed and purified from *E. coli* allowing the steps of the RK-682 biosynthetic pathway to be reconstituted *in vitro* (Sun et al., 2010).

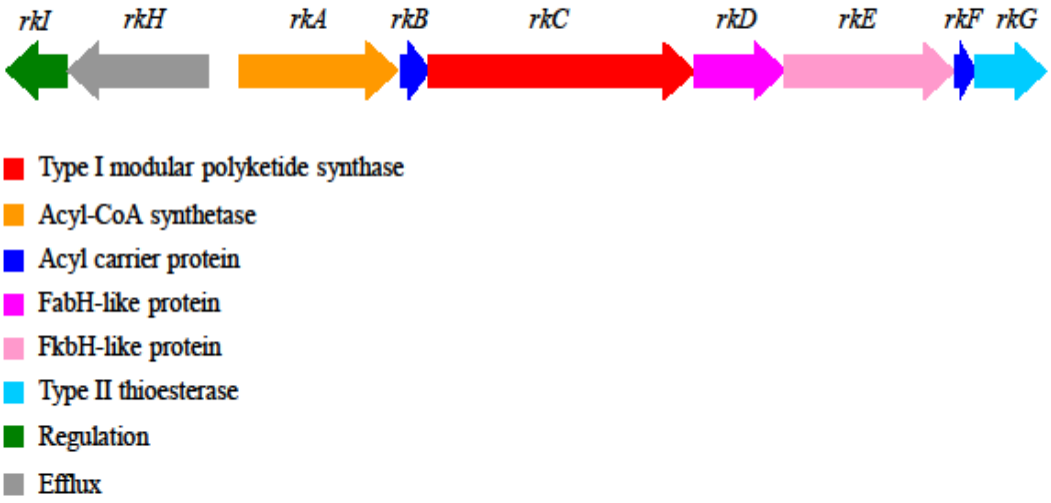


Figure 1.41 RK-682 biosynthetic gene cluster in *Streptomyces* sp. 88-682 (Sun et al., 2010).

RkA, which is a putative acyl-CoA synthetase, would activate palmitic acid and load it onto the stand-alone ACP, RkB, forming a palmitoyl-S-RkB complex (Figure 1.42). The chain would then be elongated by the PKS, RkC, which catalyses the incorporation of one malonyl-CoA unit, resulting in the formation of 3-oxo-stearoyl-S-RkC. In parallel, RkE would catalyse both the transfer of D-1,3-bisphosphoglycerate to RkF, a second stand-alone ACP, and its subsequent dephosphorylation, to form glyceryl-S-RkF. RkD, which is a putative 3-oxoacyl-ACP synthase (FabH-like), would then catalyse the condensation of glyceryl-*S*-RkF and 3-oxo-stearoyl-*S*-RkC to yield RK-682. RkD is proposed to to catalyse the C-C and C-O tetronate bond formation in RK-682 (Sun et al., 2010).

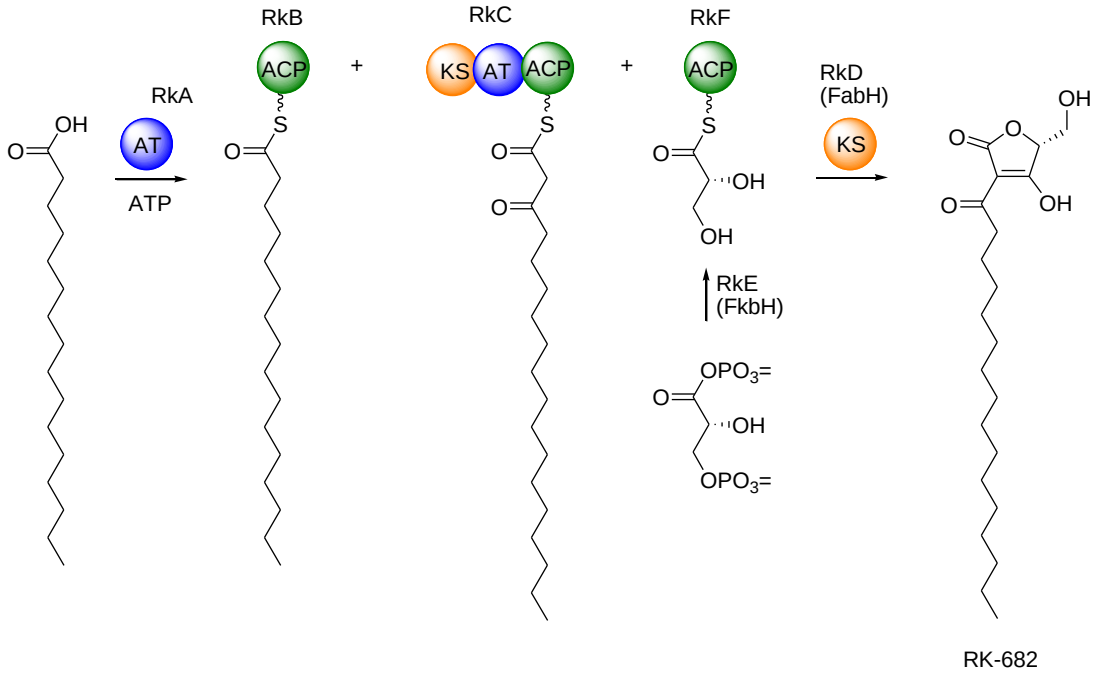


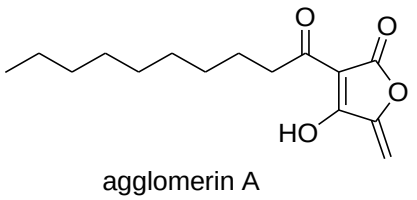
Figure 1.42 Proposed biosynthetic pathway of RK-682 (Sun et al., 2010).

1.7.2 Agglomerins

Agglomerins A (Figure 1.43), B, C and D were originally extracted from the culture broth of a bacterial strain, identified as *Pantoea agglomerans* PB-6042 (formerly *Enterobacter agglomerans*), in the late 80's by Shoji and co-workers (Shoji et al 1989). *P. agglomerans* belong to a family of Gram-negative aerobic bacteria, known as *Enterobacteriaceae*. This strain can either be commensal or pathogenic to plants and is commonly found in diverse environments, including water, soil or sediments (Cruz et al., 2007; Shoji et al., 1989). Several strains of *P. agglomerans* are used comercially in the United States as biological control agents, against the fire blight pathogen (*Erwinia amylovora*) on apple trees (Wodzinski et al., 1994; Johnson and Stockwell et al., 1998). This species is currently also considered as an opportunistic

human pathogen due to several clinical reports of local or systemic infections in immunocompromised patients (Cruz et al., 2007).

Analysis of the biological properties of agglomerins indicated its activity is predominantly against a broad range of anaerobic Gram-positive and Gram-negative bacteria, for example *Clostridium difficile, Bifidobacterium bifidum, Eubacterium aerofaciens*, and to a lesser extent against aerobic Gram-positive bacteria, for example *Streptococcus pneumoniae* and *Staphylococcus aureus* (Shoji et al., 1989). Agglomerins are structurally similar to RK-682.



agglomerin A

Figure 1.43 Structure of agglomerin A (Shoji et al., 1989).

1.8.3 Cryptic Tetronate Biosynthetic Gene Cluster in *S. scabies* 87.22.

Analysis of the *S. scabies* 87.22 genome revealed that it encodes a number of secondary metabolite biosynthetic gene clusters (Table 1.2; Yaxley, 2009). The products of some of these secondary metabolite biosynthetic gene clusters are currently unknown; such clusters are referred to as 'cryptic' biosynthetic gene clusters. One of the cryptic secondary metabolite biosynthetic gene clusters identified in *S. scabies* 87.22 (Figure 1.44) is predicted to encode proteins for the biosynthesis of a novel tetronate-like secondary metabolite similar to RK-682 from *Streptomyces sp.* 88-682 and to agglomerin antibiotics from *Pantoea agglomerans*

(Figure 1.43). It encodes genes with high sequence similarity to those present in the RK-682 biosynthetic gene cluster, such as a type I PKS, an acyl-CoA synthetase/dehydrogenase, an ACP, a FabH-like protein, and an FkbH-like protein.

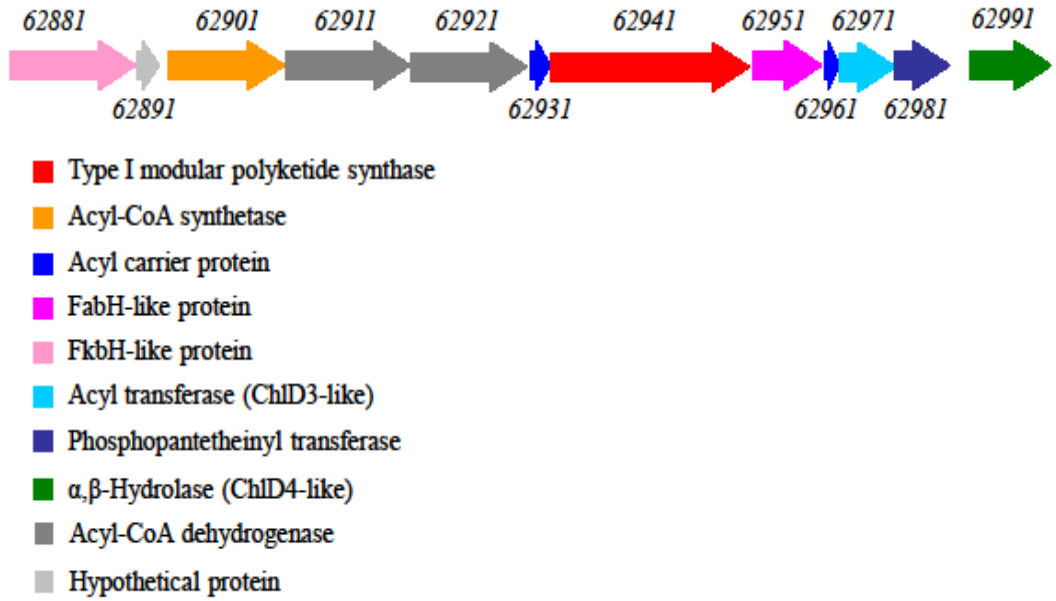


Figure 1.44 Cryptic polyketide tetronate biosynthetic gene cluster in *S. scabies* 87.22 (Challis, unpublished).

Table 1.6 Putative proteins function encoded by the putative tetronate biosynthetic gene cluster in *S. scabies* 87.22.

Protein	**Putative function**	**Homolog; origin**	**Identity (%)**
scab62881	FkbH-like protein	RkE; *Streptomyces sp.*88-682	66.1
scab62891	Unknown	-	-
scab62901	Acyl-CoA synthetase	RkA; *Streptomyces sp.*88-682	66.4
scab62911	Acyl-CoA dehydrogenase	Acyl-CoA dehydrogenase, *S. clavuligerus* ATCC 27064	83
scab62921	Acyl-CoA dehydrogenase	Acyl-CoA dehydrogenase, *S. clavuligerus* ATCC 27064	85
scab62931	Acyl carrier protein	RkB; *Streptomyces sp.*88-682	63.9
scab62941	Type I modular PKS	RkC; *Streptomyces sp.*88-682	67.9
scab62951	FabH-like protein	RkD; *Streptomyces sp.*88-682	71.7
scab62961	Acyl carrier protein	RkF; *Streptomyces sp.*88-682	64.8
scab62971	Acyl transferase (ChlD3-like protein)	Acyl transferase, *S. clavuligerus* ATCC 27064	82
scab62981	PPtase	PPtase, *S. clavuligerus* ATCC 27064	83
scab62991	α,β-hydrolase (ChlD4-like protein)	α,β-hydrolase, *S. clavuligerus* ATCC 27064	86

Based on the sequence analysis of the RK-682 biosynthetic gene cluster and the *S. scabies* 87.22 putative tetronate biosynthetic gene cluster it is likely that the *S. scabies* tetronate metabolite is assembled in a way analogous to RK-682.

The starting point of the *S. scabies* tetronate assembly would be the activation of an acyl group by the putative acyl-CoA synthetase, Scab62901, followed by loading onto the ACP, Scab62961, to form an acyl-S-ACP intermediate (Figure 1.44). The chain would then be elongated by the putative PKS, Scab62941, which would result in the formation of a β-ketothioester. In parallel, an FkbH-like protein, Scab62881, would catalyse the transfer of 3-bisphosphoglycerate to the ACP, Scab62931 and dephosphorylation by the same enzyme yields the glyceryl-S-ACP intermediate. Scab62951, which is a FabH-like protein, would then catalyse the condensation of the glyceryl-*S*-ACP intermediate with the polyketide chain attached to Scab62941, to form the tetronate ring. A further two steps are likely to be involved in *S. scabies* tetronate product formation, which are not present in RK-682 biosynthesis. These are reactions catalysed by a putative acyl-transferase, Scab62971, (ChlD3-like), and by a putative hydrolase, Scab62991, (ChlD4-like).

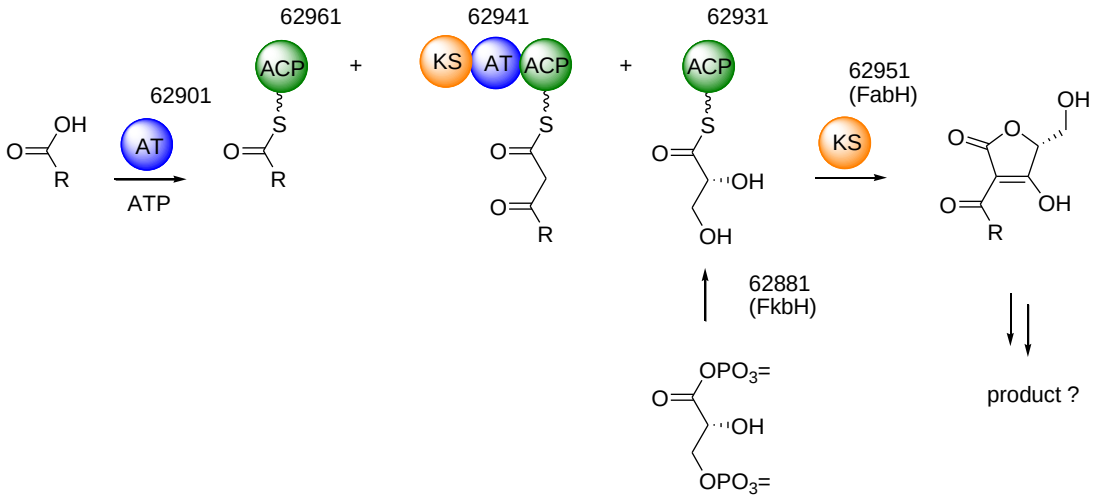


Figure 1.45 Proposed biosynthetic pathway for cryptic tetronate product in *S. scabies* 87.22 (Challis, unpublished).

1.8 Aims of the Project

The aim of this project is to investigate cryptic secondary metabolite biosynthetic gene clusters in *S. scabies* 87.22.

S. scabies 87.22 has become a model organism for studying plant-microbe interactions among Gram-positive plant pathogenic bacteria and the genome sequence for this strain is now published (GenBank Accession no. FN554889). Three out of four of the biosynthetic gene clusters investigated in this work are responsible for the production of the siderophores pyochelin, scabichelin and desferrioxamines, which are iron-chelating molecules. Iron acquisition through the use of siderophores is crucial for the development of infection in the mammalian host by pathogenic bacteria (Miethke and Marahiel, 2007). Studies of iron uptake were mostly conducted in Gram negative bacteria, whereas investigations into iron acquisition by Gram positive plant pathogens remained limited. Analysis of the *S. scabies* secondary metabolite biosynthetic gene clusters and its metabolic products is important to increase the understanding of its potential role in plant pathogenicity. For example, several genes encoding for the biosynthesis of plant toxins in these *S. scabies* secondary metabolite biosynthetic gene clusters, were found to be significant virulence factors. Awareness of these virulence factors can be used in the future in the agrochemical development to prevent infection by that pathogen potentially saving important agricultural crops from damage.

The first objective of this project was to confirm that pyochelin is the product of the putative pyochelin biosynthetic gene cluster in *S. scabies*. Analysis of the pyochelin

gene cluster mutants was planned to determine if the gene knockouts affected the production of pyochelin either by abolishing (deletion of NRPS *scab1471*) or upregulating production (deletion of *scab1401*). Comparisons of the compound from the pyochelin overproducing mutant, Δ*scab1401*, with pyochelin and enantiopyochelin authentic standards from *P. aeruginosa* were planned.

The second objective was to investigate whether the cryptic NRPS cluster in *S. scabies* directs the production of the putative hydroxamate siderophore, by creation of a partial cryptic NRPS gene deletion (*scab85471*) using a PCR targeting strategy. The mutant strain would also be examined for the production of scabichelin by comparisons of the *S. scabies* NRPS product with an authentic standard from *S. antibioticus*.

The third objective was to confirm that the putative desferrioxamine biosynthetic gene cluster in *S. scabies* directs desferrioxamine production and to investigate the role of the *desC* gene in desferrioxamine biosynthesis. The aim was to create a *desC* gene deletion in *S. scabies* using PCR targeting technology and to examine desferrioxamine production in the *desC* mutant. The function of the *desC* gene in the mutant would then be restored via genetic complementation by introducing another copy of *desC* gene into the mutant genome. The complemented mutant could then be analysed for restoration of desferrioxamine production.

The final objective was to investigate the expression of the cryptic tetronate-containing polyketide biosynthetic gene cluster in *S. scabies*. The role of the SARP family of proteins encoded for within the gene cluster in the regulation of expression of the cluster was also investigated by performing reverse transcriptase PCR reactions on the total RNA isolated from wild type *S. scabies*.

Results and Discussion I: Investigation of siderophore biosynthetic gene clusters in *S. scabies*.

2. Pyochelin as a Product of Pyochelin Biosynthetic Gene Cluster in *S. scabies.*

2.1 Analysis of the Pyochelin Biosynthetic Gene Cluster in *S. scabies* 87.22.

The pyochelin biosynthetic gene cluster (PBGC) in *S. scabies* contains 12 genes (Figure 2.1). Six of these genes encode proteins (Scab1381, Scab1411, Scab1421, Scab1461, Scab1471 and Scab1481) that are proposed to be analogous in function to the pyochelin biosynthetic proteins PchABCDEFG from *P. aeruginosa* (Patel and Walsh, 2001; Quadri et al., 1999; Serino et al., 1997).

Scab1381 is proposed to encode for a putative salicylate synthase and it contains an anthranilate synthase domain. Scab1411 is predicted to be AMP-ligase since it contains putative active site residues that are highly similar to the active site residues of MbtA, an AMP-ligase from *Mycobacterium tuberculosis*, which adenylates salicylate in mycobactin biosynthesis (Challis et al., 2000; Quadri et al., 1998; Rausch et al., 2005). Two genes *scab1481* and *scab1471* encode proteins with putative functions analogous to PchF and PchE, respectively, both of which are nonribosomal peptide synthetase (NRPS) enzymes in *P. aeruginosa* (Figure 2.1). Scab1481, being similar to PchE, is proposed to catalyse assembly of a 2-hydroxyphenylthiazolinyl thioester from the salicyl adenylate and Scab1471, analogously to PchF, adds a second molecule of cysteine to the thioester to form 2-hydroxyphenylbis-thizaolinyl thioester attached to the T domain of Scab1471 (Figure 2.2). *Scab1461* encodes a reductase enzyme homologous to PchG and it is predicted

to catalyse the reduction of one of the thiazoline rings in the thioester. Scab1471 catalyses N-methylation of the reduced thiazoline ring, followed by the release of pyochelin by its N-terminal thioesterase domain (Patel and Walsh, 2001). *Scab1421* gene is proposed to encode for a putative thioesterase analogous to PchC. PchC can catalyse removal of wrongly charged molecules from the thiolation domains of PchE and PchF (Reinmann et al., 2004).

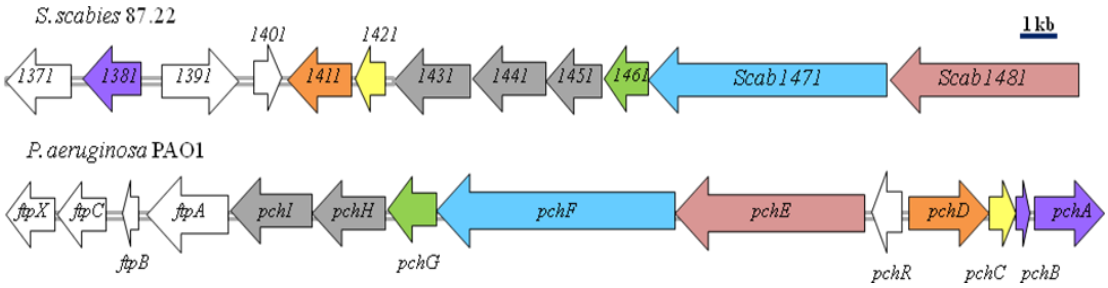


Figure 2.1 Organization of the PBGCs in *S. scabies* 87.22 and *P. aeruginosa* PAO1.

Other genes from the putative pyochelin gene cluster include *scab1401*, which encodes a pathway specific regulator homologous to PchR, a pyochelin-responsive repressor protein in *P. aeruginosa*. Scab1401 belongs to the TetR family of transcriptional regulators. Analysis of the transcription of PBGC revealed that Scab1371 is a positive transcriptional regulator of the cluster and it is co-transcribed with the predicted salicylate synthase gene *scab1381*. *Scab1431*, *scab1441* and *scab1451* encode putative transport proteins. Proposed functions of the proteins encoded by the *S. scabies* 87.22 pyochelin biosynthetic gene cluster are displayed in the Table 1.7 (Seipke et al., 2011).

Table 1.7 Putative proteins function encoded by the putative pyochelin biosynthetic gene cluster in *S. scabies* 87.22 (adapted from Seipke et al., 2011).

Protein	**Putative function**	**Homolog; origin**	**Identity (%)**
Scab1371	AfsR-family regulator	AfsR; *S. coelicolor*	40
Scab1381	Salicylate synthase	TrpE; *Saccharopolyspora erythraea*	51
Scab1391	Unknown	Sav760; *S. avermitilis*	72
Scab1401	TetR-family regulator	OrfH2; *S. griseoruber*	55
Scab1411	Salicyl-AMP-ligase	DhbE; *B. subtilis*	50
Scab1421	Thioesterase	PchC; *P. fluorescens*	44
Scab1431	ABC transporter permease/ ATPase	Nfa28020; *Nocardia farcinica*	66
Scab1441	ABC transporter permease/ ATPase	Nfa28010; *N. farcinica*	65
Scab1451	Major facilitator superfamily protein	Fraal6304; *Frankia alni*	46
Scab1461	Thiazoline reductase	Strop2818; *Salinispora tropica*	49
Scab1471	Nonribosomal peptide synthetase	PchF; *P. fluorescens*	45
Scab1481	Nonribosomal peptide synthetase	Pst2602; *P. syringae*	40

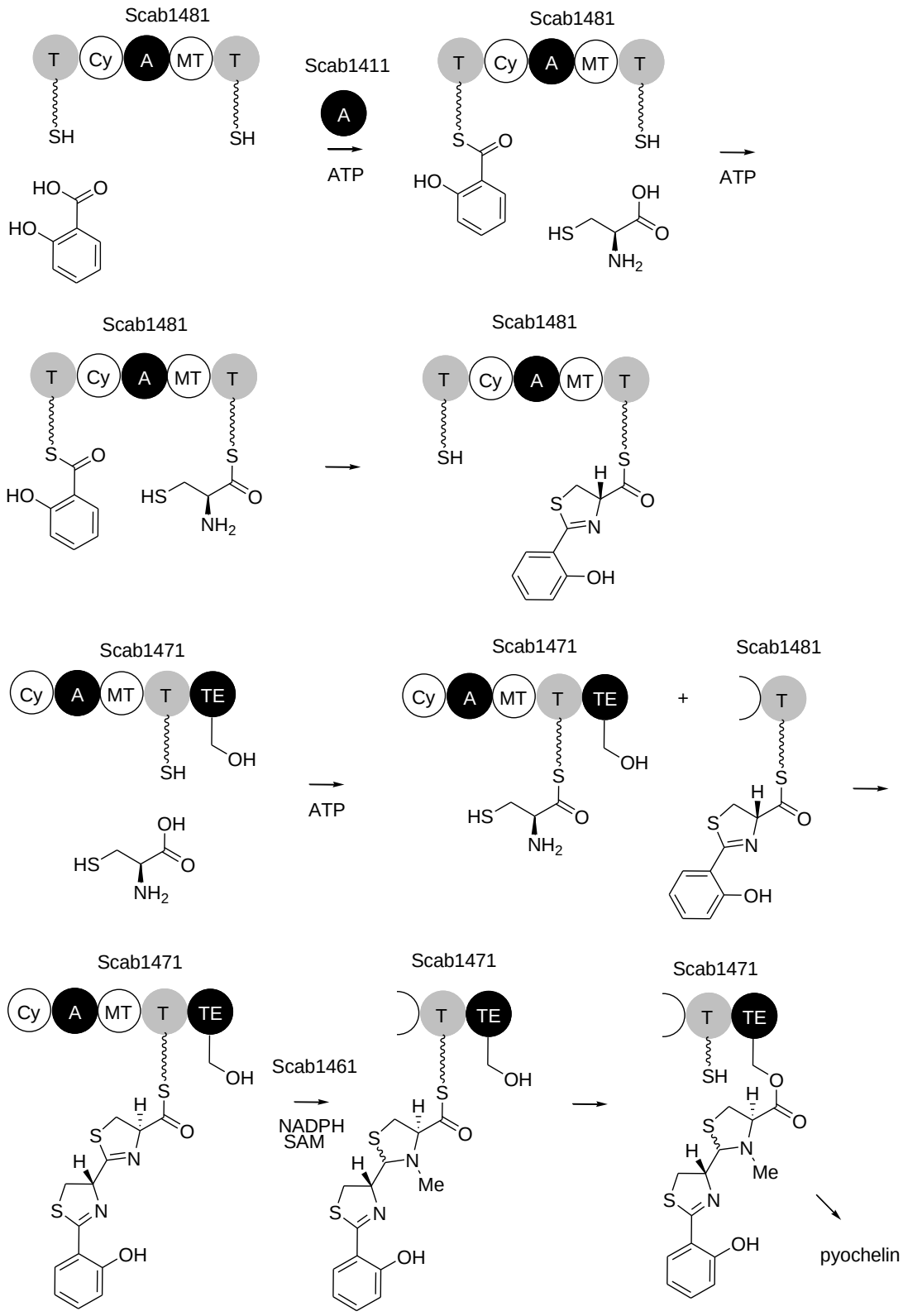


Figure 2.2 A proposed pathway for pyochelin biosynthesis in *S. scabies* 87.22 based on analysis of biosynthesis by orthologous proteins from *P. aeruginosa* (Seipke et al., 2011).

2.2 Creation of Pyochelin Mutants in *S. scabies* 87.22.

Mutants of *S. scabies* 87.22 were created using PCR-targeting REDIRECT technology (Gust et al., 2003) in the lab of Rose Loria (Cornell University). As discussed in section 2.1 *scab1401* is a gene encoded for in the putative pyochelin biosynthetic gene cluster (PBGC) in *S. scabies* 87.22 wild type strain. Analysis of pyochelin production in the wild type strain and in the Δ*scab1401* mutant revealed that Scab1401 is a transcriptional repressor of the *S. scabies* PBGC (Seipke et al., 2011). The Δ*scab1401* mutant contains a cassette consisting of *oriT* and an apramycin resistance gene from pIJ773 in place of the *scab1401* gene (Gust et al., 2003) Deletion of *scab1401* induces expression of the putative pyochelin biosynthetic gene cluster (confirmed by RT-PCR analysis in Loria lab; Figure 2.3; Seipke et al., 2011).

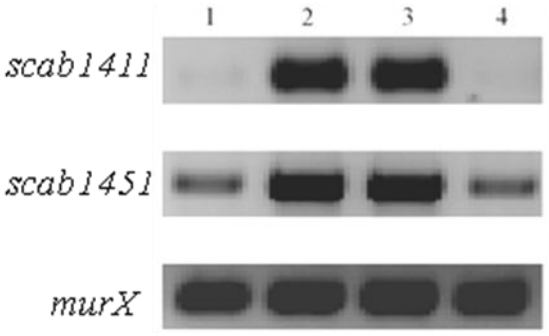


Figure 2.3 RT-PCR analysis of *ΔScab1401* mutant strain showing that Scab1401 represses transcription of *S. scabies* 87.22 pyochelin biosynthetic genes. Lane 1- *S. scabies* wild-type cDNA, 2- Δ*Scab1401*, 3- Δ*Scab1401* with pAU3-45 or 4- pRFSRL34 (plasmids containing the copy of *scab1401* gene) grown in liquid minimal medium+100 mM FeCl3 was used as template for PCRs. The murX gene -a loading control (Seipke et al., 2011).

scab1471 encodes a putative NRPS multienzyme complex homologous to *PchF* in the *P. aeruginosa* PBGC, involved in catalysing the assembly of pyochelin. The Δ*scab1471* mutant contains a cassette with *oriT* and a hygromycin B resistance gene from pIJ10700 plasmid, in place of *scab1471* (Gust et al., 2003). Deletion of *scab1471* was created to investigate the effect on pyochelin production in *S. scabies*. *scab1471* was also replaced with a gene conferring resistance to hygromycin B in the Δ*scab1401* mutant, creating the double mutant Δ*scab1401*-Δ*scab1471* by the group of Rose Loria. The creation of the double mutant was required for investigations into the function of Scab1471 NRPS under conditions giving constitutive expression of the *S. scabies* PBGC.

2.3 Analysis of Pyochelin Production in *S. scabies*

The production of pyochelin in *S. scabies* wild type and mutant strains Δ*scab1401*, Δ*scab1471* and Δ*scab1401*-Δ*scab1471* was examined by growing the strains in a liquid iron deficient medium (IDM) and then analysing the culture supernatants for the presence of pyochelin by LC-MS. Iron deficient media is a liquid minimal media in which all traces of iron are removed by Chelex resin (section 7.1.6.1; Barona-Gomez et al., 2004).

LC-MS analyses of the *Δscab1401* mutant supernatant unveiled the presence of a compound with *m/z* of 325.0669 and a retention time of 23.5 min (Figure 2.4 A). This *m/z* correlates with the *m/z* for the $[M+H]^+$ ion of pyochelin. This compound from the Δ*scab1401* mutant culture supernatant was also compared with the

simulated mass spectrum of pyochelin, with the molecular formula of $C_{14}H_{17}N_2O_3S_2^+$. It indicated that the *m/z* for each isotopic isomer of *S. scabies* compound correlates and agrees with the predicted *m/z* values for pyochelin. The relative abundances of these isotopic isomers in both the measured and the simulated spectra also agree with each other.

LC-MS analysis of the Δ*scab1471* mutant and of Δ*scab1401-1471* double mutant culture supernatants revealed that pyochelin was not produced (Figure 2.5).

The stereochemistry of the compound produced by *S. scabies* was also investigated. Research recently unveiled that the plant pathogen *P. fluorescens*, which is a Gram-negative bacterium, produces diastereoisomers of pyochelin, referred to as enantiopyochelins (ent-pyochelin I and II; Figure 2.4 C; Youard et al., 2007). Comparison of the *Δscab1401* mutant culture supernatant by LC-MS using chiral stationary phase with the synthetic standards of pyochelin and ent-pyochelin (standards kindly provided by Cornelia Reinmann) proved that *S. scabies* produce pyochelin, not enantiopyochelin (Figure 2.4 B).

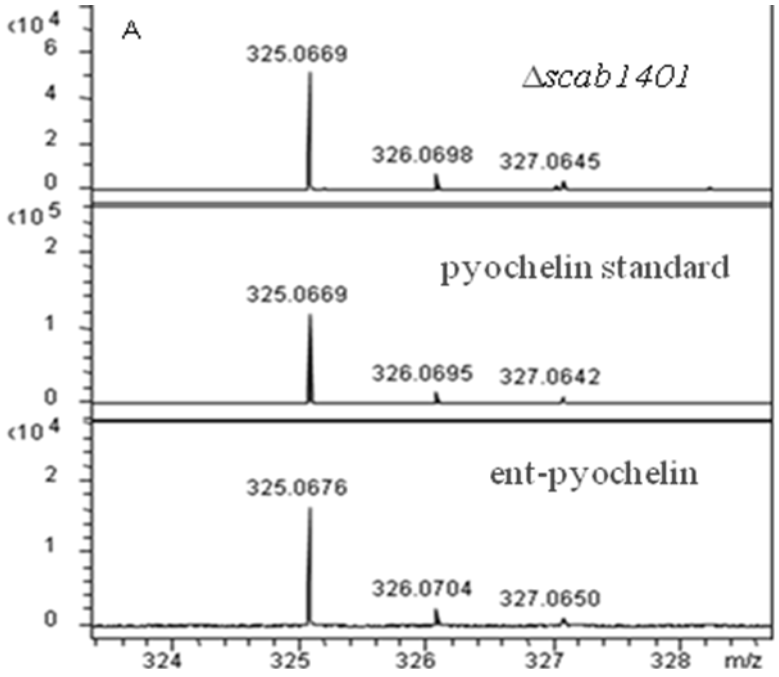


Figure 2.4 A- Measured mass spectrum of pyochelin from Δ*scab1401* supernatant (top panel), pyochelin standard (middle panel), and ent-pyochelin (bottom panel).

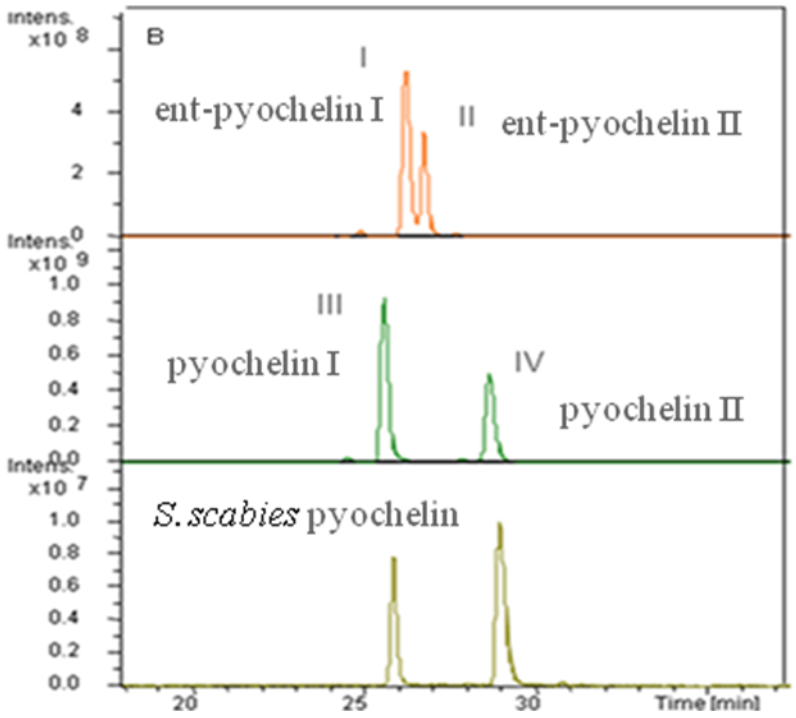


Figure 2.4 B- Extracted Ion Chromatograms (EICs) at m/z=325 from analyses of ent-pyochelin standard (top trace), pyochelin standard (middle trace) and *S. scabies* pyochelin (bottom trace).

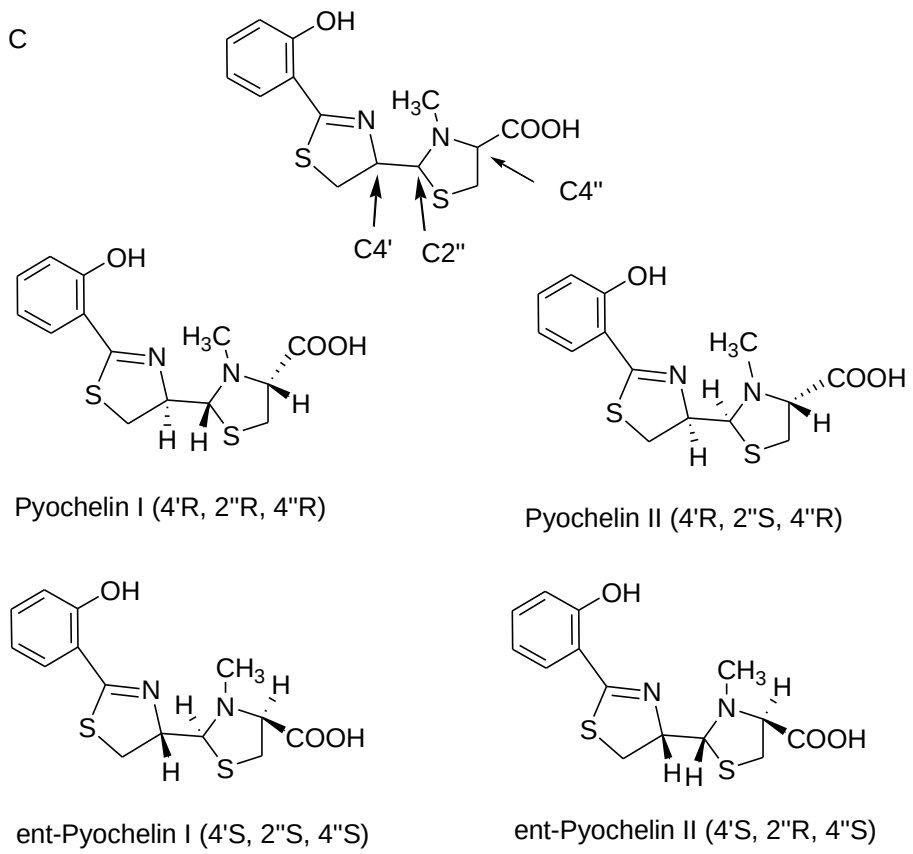


Figure 2.4 C- Pyochelin and ent-pyochelin stereoisomers. Metal induced shift at C-2'' results in isomerisation of pyochelin (Ino, and Murabayashi, 2001; Schlegel et al., 2004).

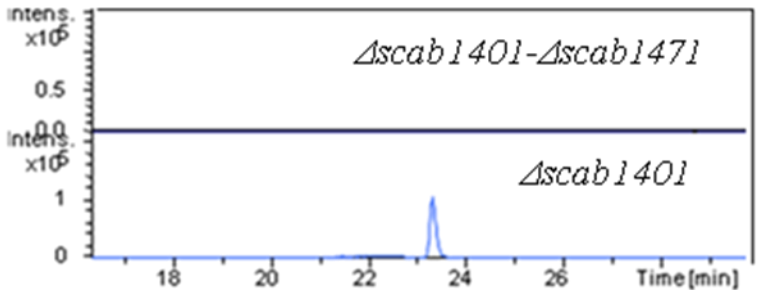


Figure 2.5 EICs at m/z= 325 from analyses of supernatants from the Δ*scab1401* mutant (bottom trace) and theΔ*scab1401*–Δ*scab1471* mutant (top trace).

2.4 Discussion

Pyochelin production by Gram-negative bacterial pathogens has been reported previously (Cox et al., 1981; Thomas, 2007; Castignetti, 1997). *S. scabies* is the first Gram-postive plant pathogen, which has been found to produce pyochelin. Mining of the *S. scabies* genome for putative secondary metabolite clusters identified a putative PBGC. Five genes from that cluster *scab1411, scab1421, scab1461, scab1471* and *scab1481* are predicted to encode proteins homologous to well characterised pyochelin biosynthetic proteins PchCDEFG from *P. aeruginosa*. The main differences between the *S. scabies* PBGC and the *P. aeruginosa* pyochelin biosynthetic gene cluster are in the organisation of the genes and in the regulation of transcription of the gene clusters.

Expression studies of the *S. scabies* PBGC, conducted by the research group of our collaborator professor Rose Loria revealed that it is transcribed as two operons. One of these is ~19 kb in size and includes the pyochelin biosynthetic genes, and the other smaller operon, which is ~3kb, consists of *scab1371* and the salicylate synthase gene *scab1381*. The two genes *scab1391* and *scab1401* are not transcribed in either operon and RT-PCR did not give any PCR products for *scab1391-scab1401* region. Scab1401 is a member of the TetR family of proteins and represses the transcription of the large biosynthetic operon during growth in high iron conditions. The other regulator Scab1371 protein is an AfsR-family regulator, which activates transcription of the large operon if the availability of iron is low (Seipke et al., 2011). In contrast, the pyochelin cluster in *P. aeruginosa* is regulated by the AraC-family protein PchR

and the ferric uptake protein, Fur (Heinrichs and Poole, 1993, 1996; Michel et al., 2005).

Pyochelin biosynthesis differs slightly between *S. scabies* and *P. aeruginaosa*. The synthesis of salicylate, a substrate for pyochelin biosynthesis in the *S. scabies* PBGC is thought to be catalysed by one protein Scab1381, while in *P. aeruginosa* the biosynthesis of salicylate is catalysed by two proteins PchA and PchB (Serino et al., 1995).

LC-MS analysis of the Δ*scab1401* mutant indicated the presence of a compound with an *m/z* value corresponding to that of pyochelin. The compound with the same *m/z* was not consistently detected in the supernatants of wild type *S. scabies* when grown in iron deficient conditions. Pyochelin should have been produced consistently in the wild type strain in iron deficient media as the pyochelin biosynthesis genes have been found to be expressed under these growth conditions. LC-MS analyses of the Δ*scab1401-1471* mutant culture supernatant, in which the pyochelin biosynthetic gene is deleted, revealed that pyochelin is not produced, further confirming that pyochelin is the product of the putative pyochelin biosynthetic gene cluster identified in *S. scabies*. Further LC-MS analysis of the Δ*scab1401* mutant culture supernatant alongside authentic standards of pyochelin and enantiopyochelin confirmed that *S. scabies* produces pyochelin and not its enantiomer enantiopyochelin.

Plant assays were used (by Prof R. Loria's group at Cornell University; Seipke et al., 2011) to analyse the importance of pyochelin in plant pathogenicity. Incubation of the Δ*scab1471* mutant (*scab1471* codes for one of the NRPS enzymes from the pyochelin biosynthetic pathway) and of the *S. scabies* 87.22 wild type strain, both on potato tissue and radish seedlings, showed tissue necrosis by both wild type and

Δ*scab1471* mutant strains indicating that pyochelin is not required for pathogenicity of potato tissue and radish. One conclusion that can be drawn from these results is that iron may not have been essential for *S. scabies* growth in these assays. The genome of *S. scabies* possesses several more biosynthetic gene clusters coding for other siderophores, such as putative genes for the biosynthesis of desferroxamines (*scab57921–scab57951*). To determine which iron-chelating metabolites produced by *S. scabies* are required for plant pathogenicity, analysis of a desferrioxamine mutant and other siderophore mutants in *S. scabies* is needed, as well as, in planta transcriptional analysis of the pyochelin, desferrioxamine and other putative siderophore biosynthetic gene clusters.

3. Investigation of a Putative Hydroxamate Siderophore Biosynthetic Gene Cluster in *S. scabies* 87.22

3.1 PCR Targeting Strategy

Mutagenesis of *S. scabies*, to create gene knockouts, was achieved using a PCR targeting protocol (Gust et al., 2002; Gust et al., 2003). This strategy allows the replacement of a genomic DNA sequence, located in the cosmid genomic library of *S. scabies*, with a cassette containing an apramycin resistance gene. This cassette is amplified by PCR using primers, which have 39 nt 5' regions homologous to the regions adjacent to the targeted gene. The cassette also contains the origin of transfer (*oriT*) allowing for the transfer of the PCR targeted cosmid from an *E. coli* strain to *S. scabies* via conjugation.

Plasmid pIJ773 (Figure 3.1) was used for obtaining the resistance cassette. The plasmid containing an apramycin gene and *oriT* was digested using restriction enzymes *HindIII* and *EcoRI* resulting in two fragments. The fragment of 1.4 kb, containing the apramycin resistance gene and *oriT*, was purified from an agarose gel and subsesquently used as a PCR template.

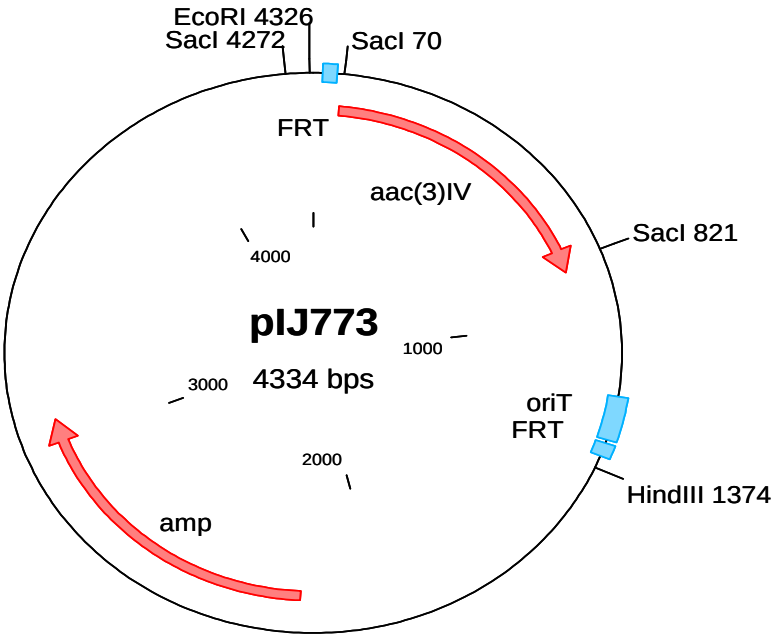


Figure 3.1 Map of pIJ773 (Gust et al., 2003).

The PCR primers for amplification of the disruption cassette were 59 nt (upstream primer) and 58 nt (downstream primer). The sequence of each set of primers had at the 5' end, 39 nt homologous to the *S. scabies* DNA region adjacent to the gene to be inactivated and at the 3' end, 19 or 20 nt same as the right or left end of the resistance cassette (Figure 3.2).

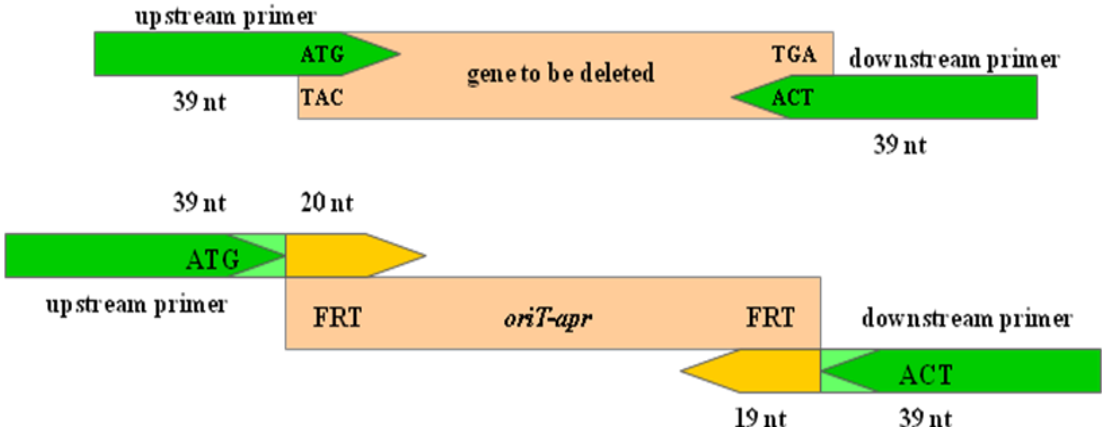


Figure 3.2 Design of PCR primers for making a gene replacement (Gust et al., 2002).

The cosmid with the gene to be disrupted was transferred into *E. coli* BW25113, which contains the recombination plasmid pIJ790. The apramycin cassette was then introduced into cosmid by electroporation. The resulting apramycin resistant colonies were ones in which a double cross over between the target sequence in the cosmid and the cassette has occurred. These colonies were subsequently analysed by restriction digest and PCR, using a pair of test primers. These test primers were 18-20 nt long and were designed to prime approximately 100 bp outside of the region of the gene disruption. The size of the resulting PCR products makes it possible to distinguish between the presence of the wild type gene in the cosmid and the disruption cassette.

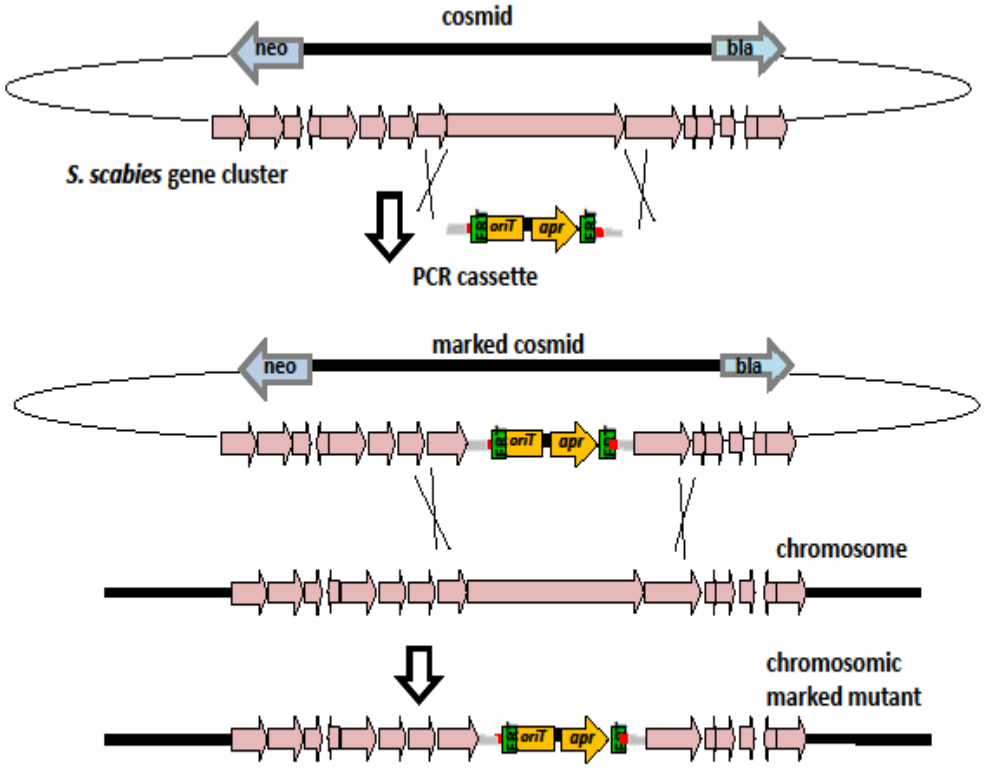


Figure 3.3 Creation of a gene deletion in the cosmid via homologous recombination of the disruption cassette with cosmid DNA and then conjugation from *E. coli* into *S. scabies* and double homologous recombination resulting in a replacement of the gene of interest with the disruption cassette creating an *S. scabies* mutant. Yellow gene – apramycin resistance, orange region – origin of transfer (*oriT*), green regions = FRT sites.

The cosmid with the replaced gene was then introduced via transformation into *E. coli* ET12567/pUZ8002, which is a DNA non-methylating strain (Figure 3.3). It is necessary to use this strain as *S. scabies* carries a methylated DNA-sensing restriction system, which degrades methylated DNA from ordinary *E. coli* strains. The mutagenised cosmid was then transferred from *E. coli* ET12567/pUZ8002 to *S. scabies* via conjugation. This was completed by mixing the *E. coli* ET12567/pUZ8002 cells containing the mutagenised cosmid with *S. scabies* spores, and then plating a dilution series on SFM plates. Following an overnight incubation

the plates were overlaid with nalidixic acid, killing only the *E. coli*, and with apramycin, which selectively kills the *Streptomyces* that did not contain introduced cosmid. As a result of the conjugal transfer of cosmid, the apramycin resistant *S. scabies* colonies on the SFM plates should have the target gene replaced with the resistance cassette by double crossover. The double cross over was recognized by apramycin resistance (cassette contains apramycin) and kanamycin sensitivity (the cosmid backbone containing kanamycin resistance gene has been lost during cross over). The replacement of the gene was confirmed by PCR and Southern blot hybridization.

3.2 Investigation of Putative Hydroxamate Siderophore Production in *S. scabies*

The aim of this work was to examine whether the putative hydroxamate siderophore (Figure 3.4) is present in culture supernatants of *S. scabies* 87.22 grown under iron deficient conditions. It has been reported by our collaborators Kodani and co-workers that the siderophore was isolated from the culture supernatants of *S. antibioticus* NBRC 13838 (the strains that they used in their work). That siderophore is also proposed to be the product of a cryptic NRPS enzyme encoded by *scab85471*, discovered by genome mining of *S. scabies* 87.22.

3.2.1 Isolation and Structure Elucidation of the Hydroxamate Siderophore

Analysis of the culture supernatants of *S. antibioticus* NBRC 13838, via HPLC by our collaborators, Dr Shinya Kodani and his co-workers, indicated that this strain produces a putative siderophore that was named scabichelin. Our collaborators also carried out the structural characterisation of scabichelin. The metabolite was converted into a complex with Galium before the NMR analysis, which helped signal dispersion and aided stereochemical analysis. The Ga-scabichelin complex structure was elucidated by H NMR AND 2D NMR. The NMR spectra analysis indicated that the compound consists of the five amino acids: N-methylated hydroxy-formyl-ornithine, serine, N-methylated hydroxy-ornithine, ornithine and cyclic hydroxyl-ornithine. The mass spectrum had a parent ion present for scabichelin at the *m/z* 648.3671, which agreed well with the calculated value of 648.3675 corresponding to $[C_{26}H_{50}N_9O_{10}]^+$. The stereochemistry of scabichelin was determined by chemical analysis using Marfey method (Marfey, 1984) and 3D modelling (Deeth, 2001) as LLLLD.

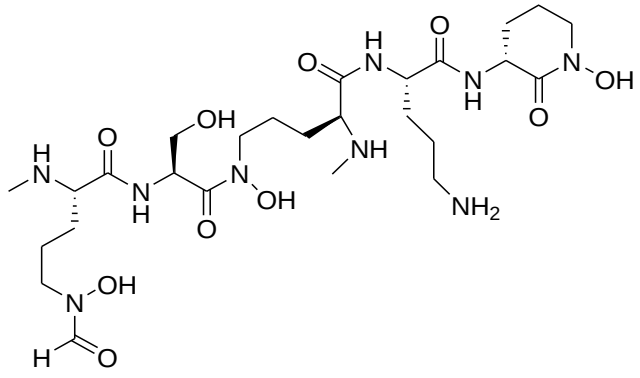


Figure 3.4 Structure of a novel siderophore scabichelin.

3.2.2 Scabichelin Biosynthetic Gene Cluster

Analysis of the *S. scabies* 87.22 genome sequence unveiled a cryptic NRPS biosynthetic gene cluster, which includes the protein coding sequences from Scab85431 to Scab85521 (Figure 3.5). All of these proteins, present within the NRPS biosynthetic gene cluster, are homologous to proteins encoded for by genes within the coelichelin biosynthetic gene cluster from *S. coelicolor* M145 (section 1.7.3; Table 3.1; http://www.ncbi.nlm.nih.gov/nuccore/260644157).

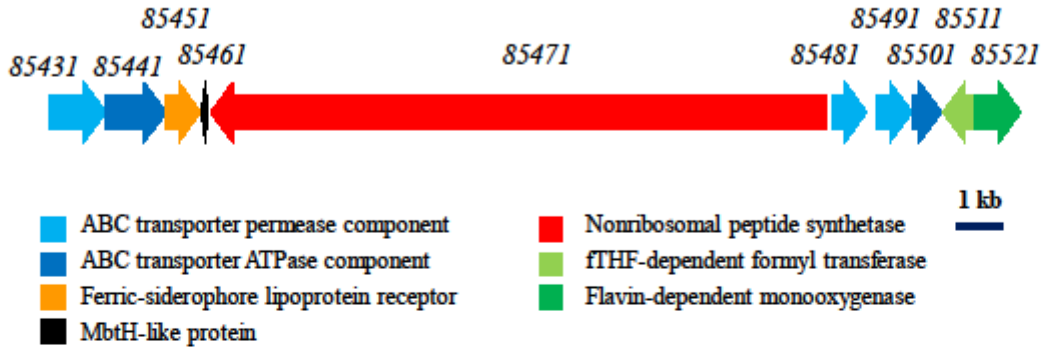


Figure 3.5 Organization of the scabichelin biosynthetic gene cluster in *S. scabies* 87.22.

Bioinformatic analysis of the functional domains of the NRPS encoded for by *scab85471* revealed it contains 17 enzymatic domains organized into 5 modules (Figure 3.6 B). Each module contains an adenylation (A), a thiolation (T) and a condensation domain (C) (http://www.nii.res.in/searchall.html). The final termination module of *scab85471* also contains an epimerization/condensation domain (Figure 3.6 B) suggesting that the stereochemistry of the peptide is L-L-L-L-D. It is hypothesised that product of the *scab85471* gene would be a pentapeptide, as it contains 5 modules.

Two genes from the putative scabichelin biosynthetic cluster, *scab85511* and *scab85521,* encode proteins homologous to CchA and CchB from *S. coelicolor* M145, respectively. These enzymes have sequence similarity to formyl transferase and NADH/FAD-dependent monoxygenases, respectively. CchB catalyses the hydroxylation of L-ornithine (L-Orn) to yield L-N5-hydroxyornithine (L-hOrn) and CchA subsequently formylates it yielding L-N5-formyl-N5-hydroxyornithine (L-fhOrn; Pohlman and Marahiel, 2008). These two non-proteinogenic amino acids are later incorporated into coelichelin. Analogous reactions are predicted to be catalysed by Scab85511 and Scab85521 in the biosynthesis of the novel hydroxamate siderophore in *S. scabies* (Figure 3.6 A).

The amino acids, which are the building blocks for the putative *S. scabies* product are loaded onto the thiolation (T) domains by the adenylation (A) domains of the NRPS encoded by *scab85471* (Figure 3.6 B).

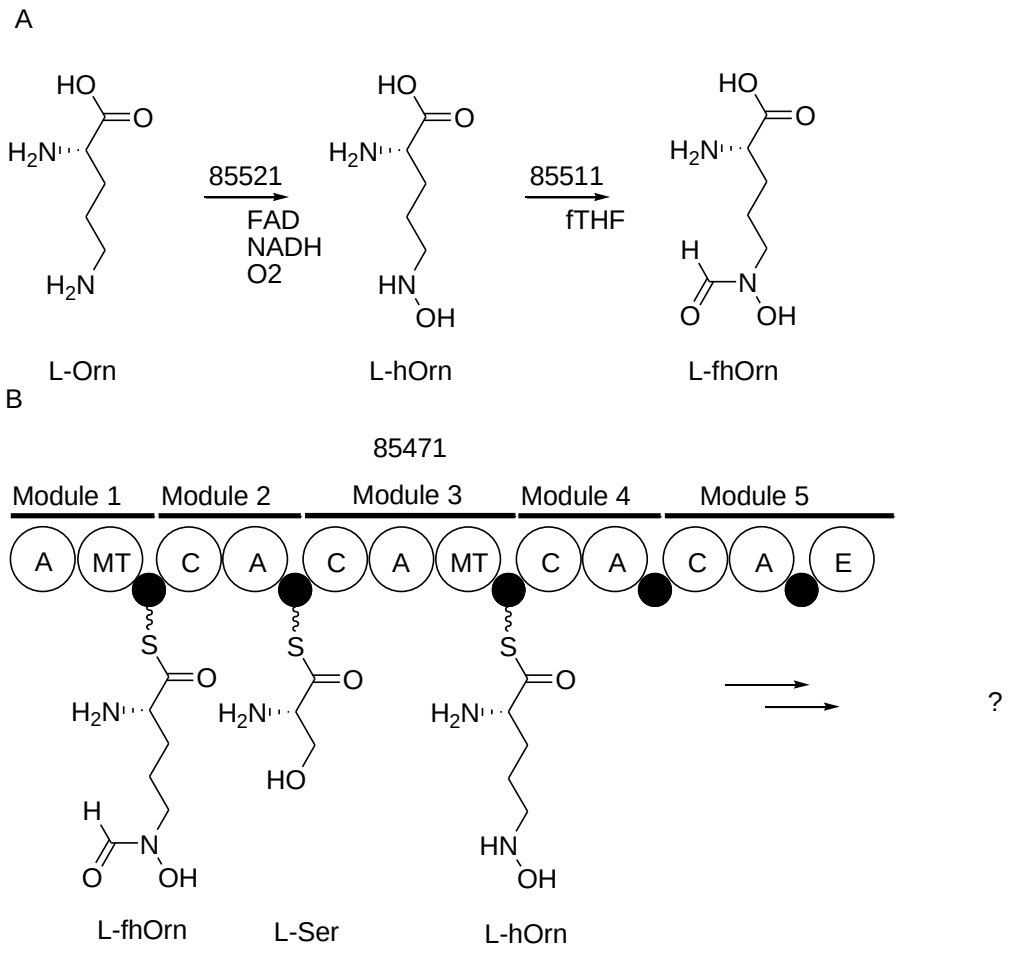


Figure 3.6 A-Proposed role of the enzymes encoded by *scab85521* and *scab85511* in the biosynthesis of the amino acids L-N5-hydroxyornithine (L-hOrn) and L-N5-formyl-N5-hydroxyornithine (L-fhOrn). B-Organisation of modules and domains of the NRPS encoded by *scab85471*. The residues attached to the thiolation domains (black circles) of module 1, 2, 3 are the aminoacids predicted to be incorporated into scabichelin by each A domain. The substrates loaded by A domains of the modules 4 and 5 could not be unambiguously predicted.

3.2.3 Creation of a *scab85471* Mutant

The *scab85471* disruption was obtained using the PCR targeting REDIRECT technology. In this method the cosmid S57 (a clone from an *S. scabies* supercos1 genomic library containing *scab85251* to *scab85521* provided by Prof R Loria of Cornell University) containing an insert with the *scab85431-scab85521* gene cluster was used. A 3 kb fragment internal to *scab85471*, which encodes an entire NRPS module 2 (Figure 3.7) was selected to be deleted, thus disrupting its function. This internal fragment of *scab85471* was chosen as this is a coding region for one of the domains of the Scab85471 NRPS. It was assumed that deletion of the sequence encoding one of the protein domains, rather than inactivation of any random region of *scab85471,* would ensure the inactivation of the target gene was achieved. The internal fragment to be deleted was chosen so that a region flanking either side of the deletion remained. Sufficient flanking sequences on both sides of target sequence would ensure that homologous recombination would occur.

Before creating the *scab85471* disruption both ends of the insert in the cosmid were sequenced to determine from which gene of the *S. scabies* biosynthetic cluster the insert starts and at which gene the insert ends. For this 20 nt forward and reverse sequencing primers in the insert were created. The primers were designed approx. 100 bp before the ends of the insert and going out from the insert towards the cosmid backbone (section 7.1.5).

The integrity of the wild type *scab85471* containing cosmid was also analysed by restriction digest with *BamH*I, *Nco*I, *Pst*I and *Not*I restriction enzymes (Figure 3.8).

The band pattern of each digest on the gel was consistent with the pattern shown by a virtual digestion program (Figure 3.9).

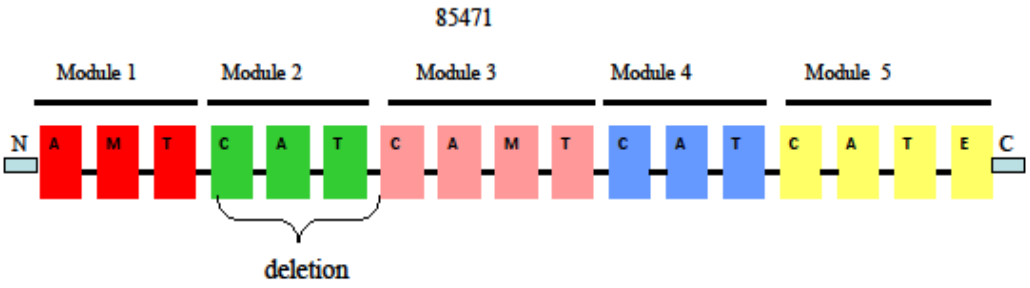


Figure 3.7 The enzymatic domains of *scab85471*. The fragment deleted is module 2.

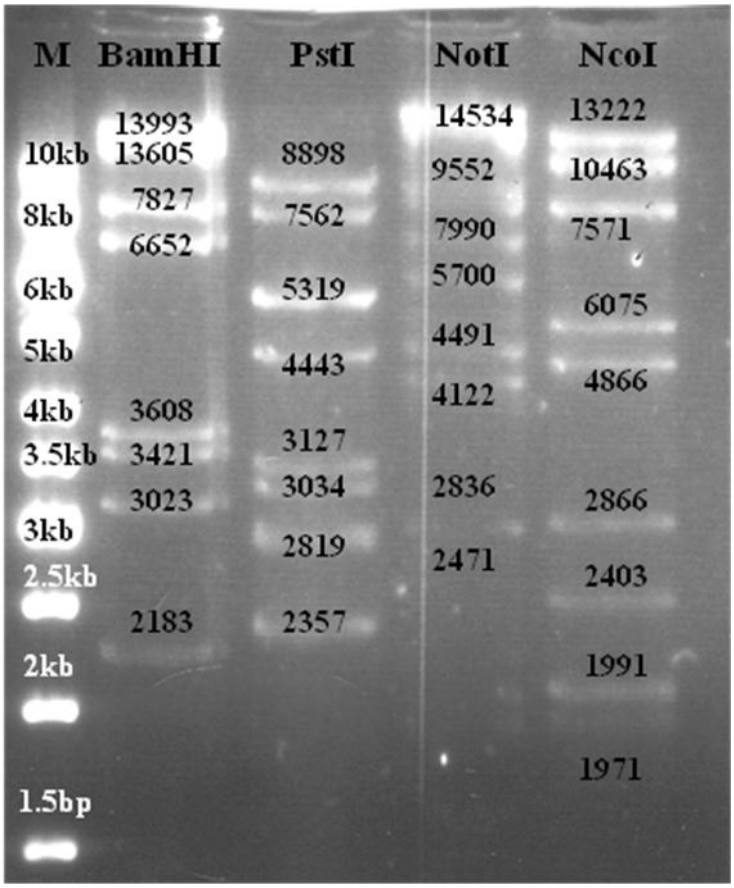


Figure 3.8 Restriction analysis of the wild type S57 cosmid. Lanes from left to right:M- 1kb DNA molecular size marker, *Bam*HI, *Pst*I, *Not*I, *Nco*I digests.

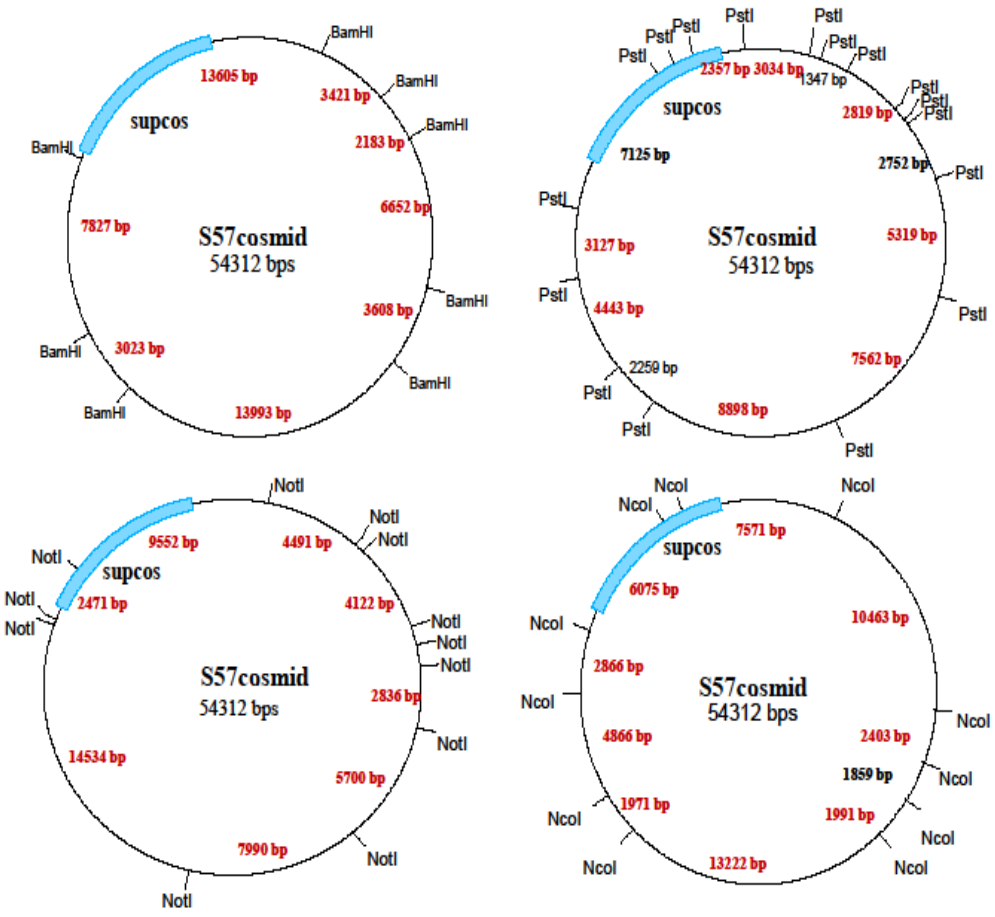


Figure 3.9 Restriction digest pattern of scabichelin cosmid (S57) with *Bam*H, *Pst*I, *Not*I and *Nco*I restriction enzymes (the band sizes in bp that are displayed red in the diagram indicate bands present on the gel picture in figure 3.8).

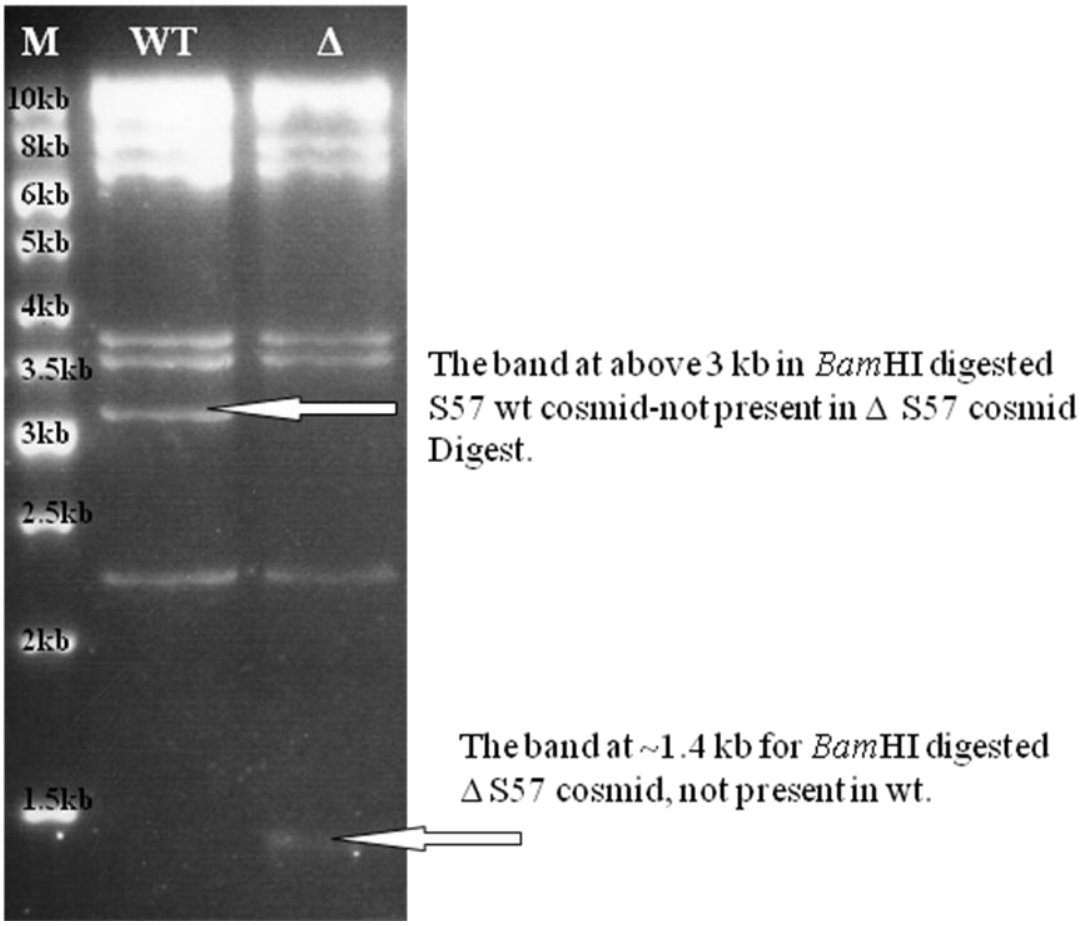


Figure 3.10 Restriction analysis of wild type vs mutagenised S57 cosmid (S57::*scab85471*::*apr-OriT*. Lanes: M- 1kb DNA molecular size marker, WT- wild type S57 cosmid *BamH*I cut, Δ- cosmid + *oriT-apramycin* cassette *BamH*I cut.

The *oriT- apramycin* cassette from the pIJ773 plasmid was amplified by PCR using primers, which have 39 nt 5' extensions homologous to the regions adjacent to the targeted internal fragment of *scab85471* (3233 bp encoding C, A and T domain of module 2). The cosmid containing *scab85471* was transferred into *E. coli* BW25113, which carries recombination plasmid pIJ790. Following this the apramycin resistance cassette was introduced into the cosmid via electroporation. The apramycin resistant colonies are ones in which the double cross over between the regions adjacent to the target sequence in the cosmid and the cassette has occurred. These colonies were

analysed by restriction digest (Figure 3.10) and PCR with a pair of test primers (Figure 3.11). These primers were 18-20 nt long (Section 7.1.5) and were designed to prime approximately 100 bp outside of the region of the gene disruption. The size of the PCR allowed for comparison between the wild type gene in the cosmid and the gene in the disruption cassette. The test primers generated the expected ~1.6 kb fragment after gene disruption (1.4kb-the size of apramycin cassette plus ~ 200bp distance of primers from the cassette and approximately 3kb for the wild type - Figure 3.11).

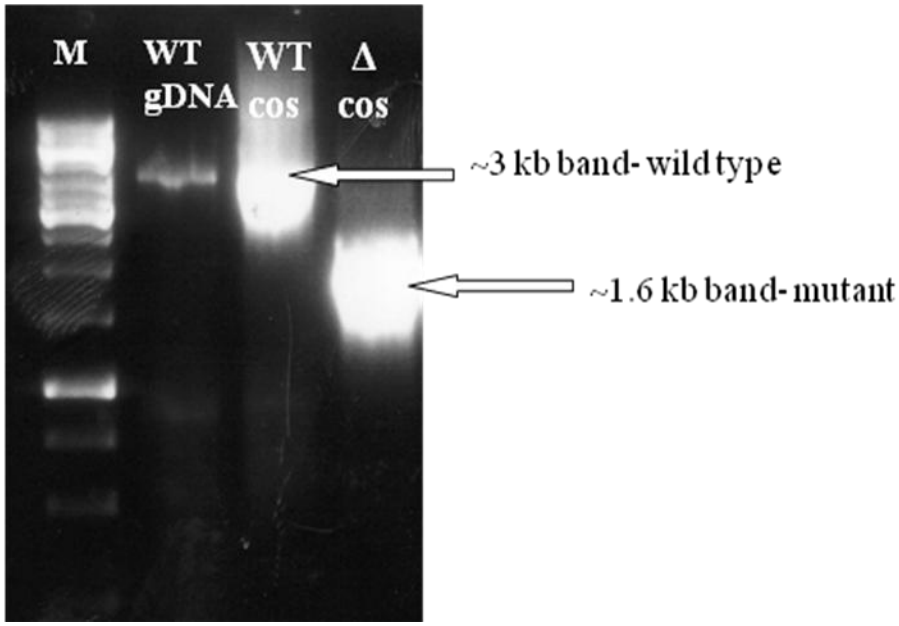


Figure 3.11 PCR analyses of the wild type and the mutagenised S57 cosmid with the test primers. Lanes: M- 1kb DNA molecular size marker, WT gDNA- PCR band for the wild type *S. scabies* DNA, WT cos- PCR band for the wild type S57cosmid, Δ cos-PCR band for the S57 cosmid + *oriT-apramycin* cassette.

The cosmid with replaced gene was introduced by transformation into *E. coli* ET12567/pUZ8002, which is a DNA non-methylating strain. The mutagenised

cosmid was then transferred from *E. coli* ET12567/pUZ8002 to *S. scabies* by conjugation where a double cross over between the homologous regions in *S. scabies* genomic DNA and *scab85471* on cosmid would occur, yielding a mutant with inactivated *scab85471*. The replacement of the gene was confirmed by PCR and Southern blot hybridization. The test primers (Test_fw and Test_rv; Table 7.6) were used to analyse replacement of the target sequence in the genome by PCR (Figure 3.12). The mutant was named *S. scabies* W1000.

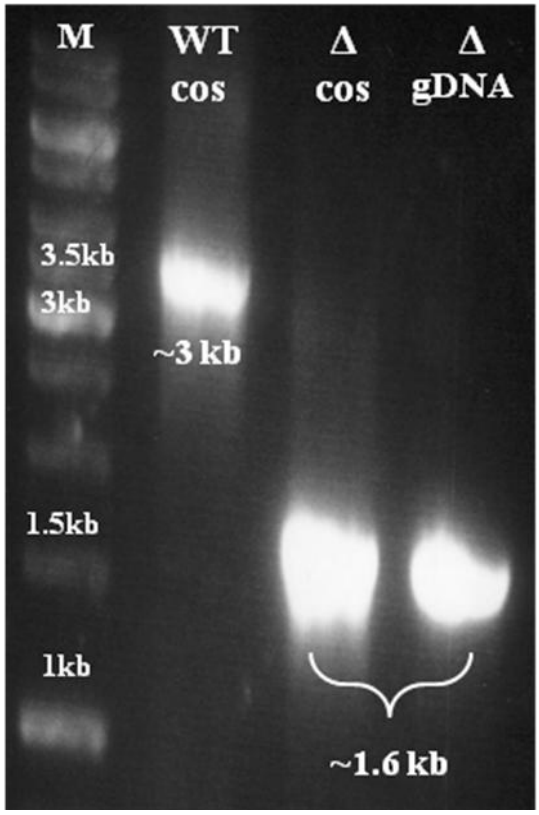


Figure 3.12 PCR analysis of the genomic DNA from the putative scabichelin mutant. Lanes: M- 1kb DNA molecular size marker, WT cos- PCR band for the wild type scabichelin cosmid, Δ cos-PCR band for the scabichelin cosmid with *OriT-apramycin* cassette, Δ gDNA- PCR for the for the scabichelin mutant genomic DNA.

Southern blot hybridization analysis was carried out to analyse the position of the resistance cassette within the *S. scabies* W1000 genome. Genomic DNA and the

cosmid containing mutagenised scabichelin gene were digested by *Bam*HI. A DIG-labelled mutagenised scabichelin cosmid was used as a probe, which was hybridised to the digested DNA. Comparisons could then be made between the wild type and mutant on an X-ray film. Bands that could be seen on the x-ray film were highlited in white, red or green (Figure 3.13).

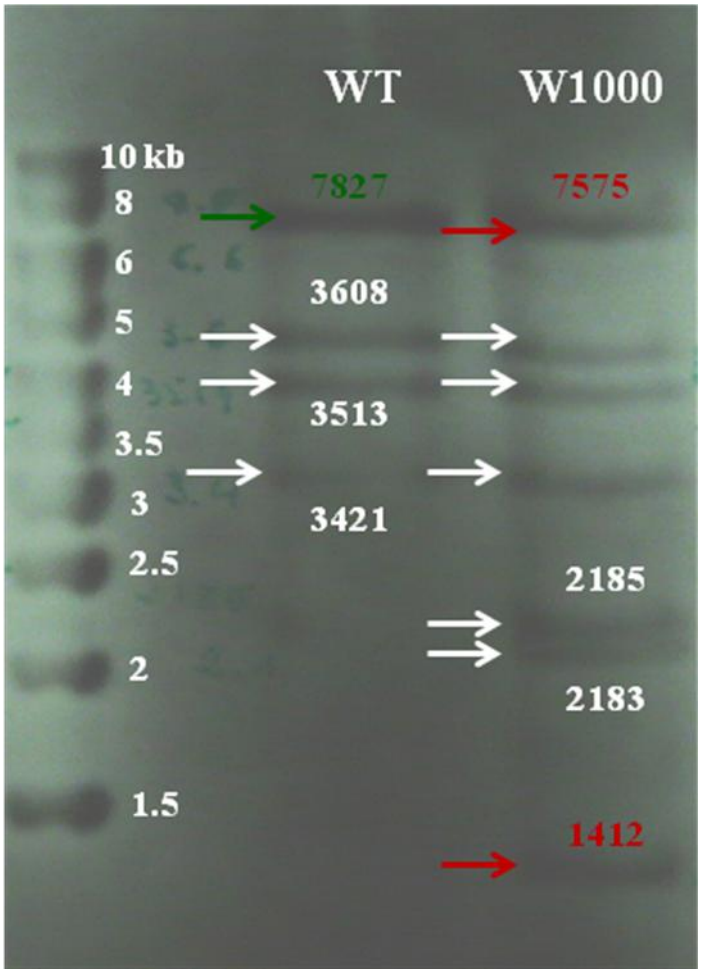


Figure 3.13 Southern blot hybridization using labelled S57::*oriT-apr* cosmid as a probe confirming the nature of *S.scabies*W100 mutant with 3 kb fragment deleted within *scab85471* gene. Bands in white colour are all present in the *S.scabies* wild type and the mutant. Band 7827 bp in green is only present in the wild type strain. Bands 1412 bp and 7575 bp coloured in red are characteristic for the mutant.

The mutant was also analysed by sequencing using Test_fw and Test_rv primers (Section 7.1.5), which confirmed the presence of the apramycin resistance gene in the *S. scabies* W1000 genome.

3.2.4 LC-MS Analysis of Scabichelin Mutant

LC-MS analysis of the W1000 scabichelin mutant culture supernatant compared with that of wild type culture supernatant, with the ferric iron added before analysis, indicated the presence of a ferric complex in *S. scabies* 87.22 but not in the W1000 mutant (Figure 3.14). This complex had identical retention time (Figure 3.14) and MS/MS fragmentation pattern (Figure 3.15) to scabichelin isolated and purified from *S. antibioticus* (provided by our collaborators), which was used as an authentic standard. These data conclusively proves that scabichelin is the metabolic product of the Scab85471 NRPS in *S. scabies* 87.22.

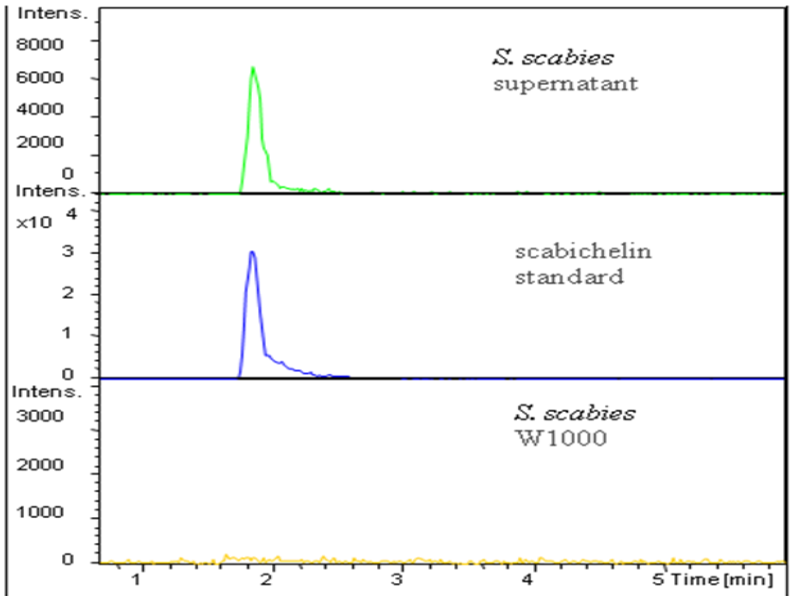


Figure 3.14 EICs at *m/z*=351.14 (corresponding to the $[M+2H]^{2+}$ ion for the ferric-scabichelin) from analyses of the culture supernatant of *S. scabies* 87.22 (top panel), Fe-scabichelin purified from *S. antibioticus* (middle panel) and the culture supernatant of *S. scabies* W1000 (bottom panel).

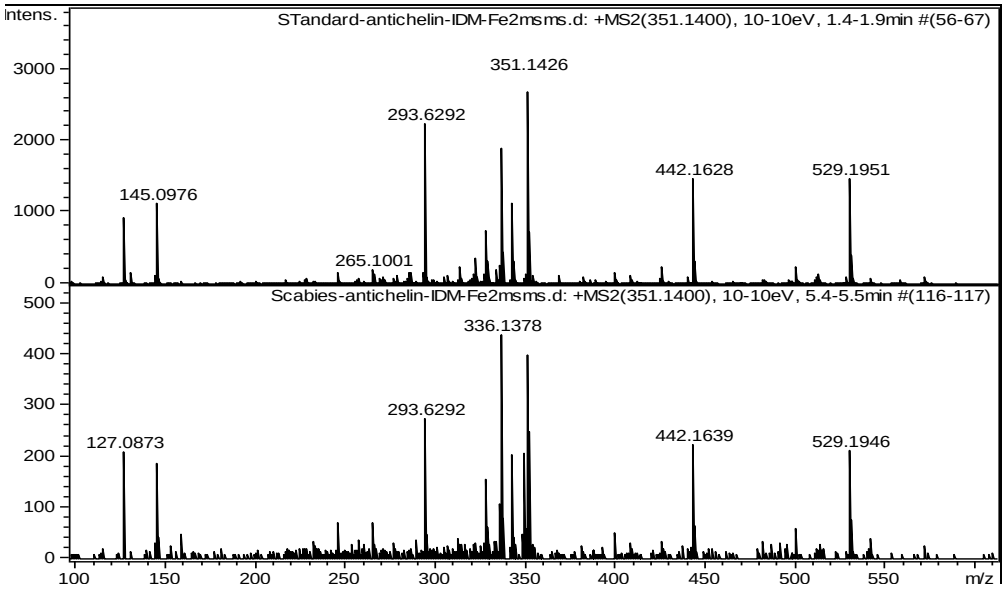


Figure 3.15 MS/MS spectra of ferric-scabichelin purified from *S. antibioticus* (top panel) and from ferrated supernatants from iron-deficient cultures of *S. scabies* 87.22 (bottom panel).

3.2.5 Discussions

Analysis of the *S. scabies* 87.22 genome sequence unveiled a cryptic NRPS biosynthetic gene cluster, which includes the protein coding sequences from *scab85431* to *scab85521* (Figure 3.17 A). All of these proteins, present within the NRPS biosynthetic gene cluster, are homologous to proteins encoded for by genes within the biosynthetic gene cluster from *S. coelicolor* M145 (Table 3.1), which is a producer of tris-hydroxamate siderophore coelichelin (Figure 3.16; Lautru et al., 2005). It was, therefore, hypothesised that the biosynthetic cluster encoded by *S. scabies* may direct the production of a hydroxamate-containing metabolite named as scabichelin (Figure 3.16).

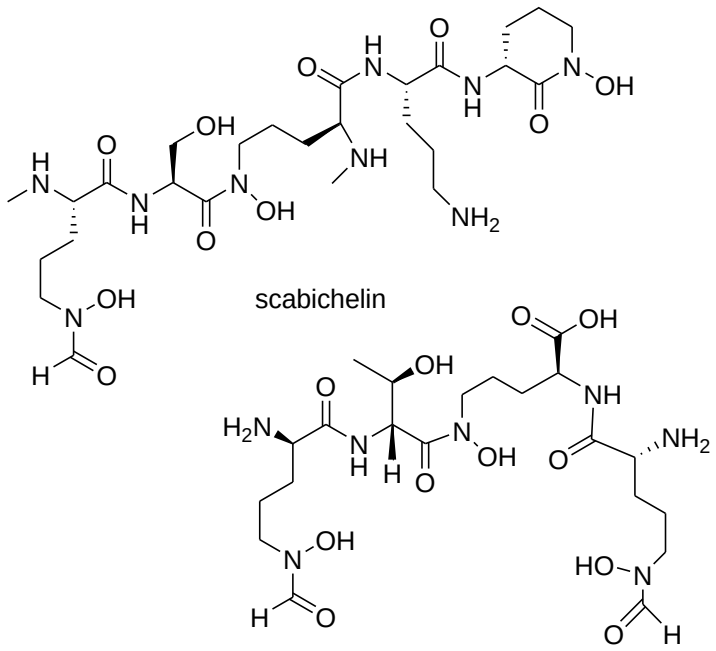


Figure 3.16 Structures of coelichelin and a novel siderophore scabichelin.

PCR targeting method (Gust et al., 2003) was used to create a mutant of *S. scabies* 87.22 in order to investigate the role of the cryptic nonribosomal peptide biosynthetic gene cluster in scabichelin biosynthesis. The mutant, named as *S. scabies* W1000, was obtained via replacement of an internal fragment of the *scab85471* gene with an *oriT-apr* cassette. LC-MS/MS analyses of culture supernatants of the 87.22 and W1000 strains grown in an iron-deficient medium, to which ferric iron was added before analysis, indicated the presence of a ferric complex that is produced by *S. scabies* 87.22, but not the W1000 mutant. LC-MS retention time and MS/MS fragmentation pattern of ferri-scabichelin are the same for *S. scabies* 87.22 compound and the authentic standard. The experimental data confirm that *scab85431-scab85521* gene cluster is involved in the production of scabichelin.

The hydroxymate siderophore coelichelin, a product of *S. coelicolor*, similarily as scabichelin, was predicted to be produced from bioinformatic analysis of *S. coelicolor* genome (Challis and Ravel, 2000). It was later isolated and structurally characterised (Lautru et al., 2005). The coelichelin biosynthetic gene cluster encodes 11 proteins (Figure 3.17 B; Challis and Ravel, 2000), of which one is a trimodular NRPS (CchH) that catalyse the condensation of, in total, four amino-acid residues to form a tetrapeptide. Structure-based sequence comparisons allowed predicting the substrates for modules 1, 2 and 3 as *N*-formyl-*N*-hydroxyornithine (L-hfOrn), L-threonine (L-Thr) and *N*-hydroxyornithine (L-hOrn), respectively (Figure 3.17 B).

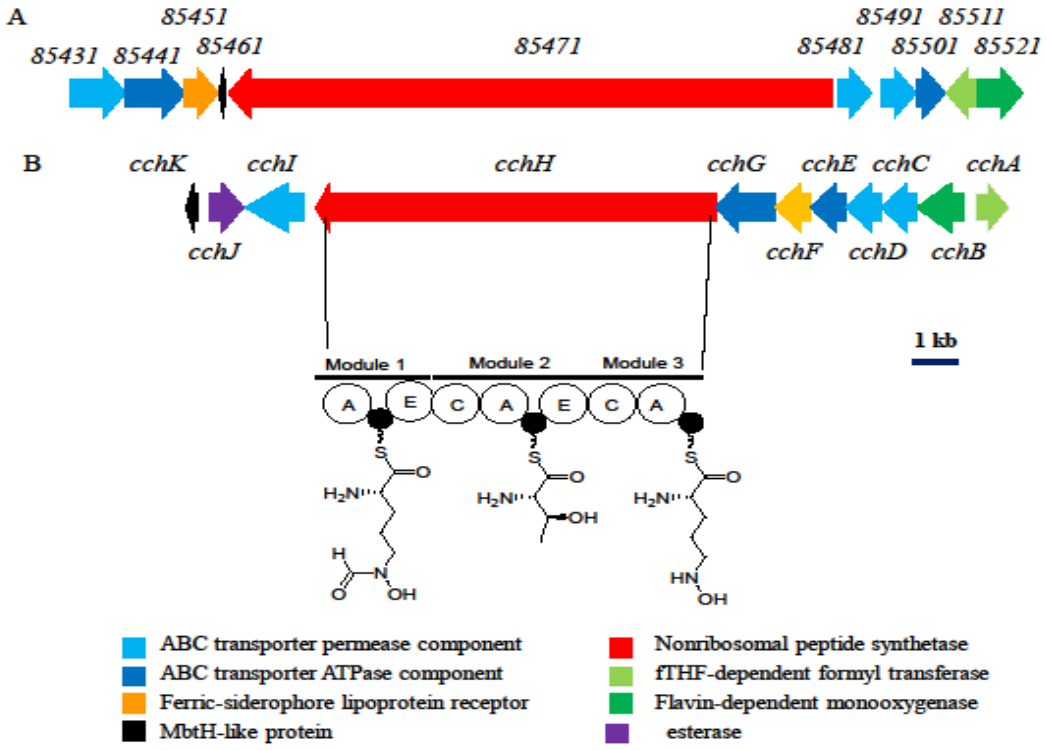


Figure 3.17 A-putative scabichelin biosynthetic gene cluster. B-Organisation of the coelichelin biosynthetic gene cluster and the module and domain organisation of the NRPS encoded by *cch*H. The residues attached to the thiolation domains (black circles) of module 1, 2, 3 are the aminoacids predicted to be incorporated into coelichelin by each A domain.

Unlike the proposed biosynthesis of scabichelin, where C domains of five modules of the NRPS are predicted to catalyse condensation of five amino acyl thioesters, coelichelin biosynthetic model (Figure 3.18) involves iterative usage of module 1 to incorporate two molecules of *N*-formyl-*N*-hydroxyornithine into the peptide (Lautru et al., 2005). In the first step, the A domains of modules 1, 2, 3 activate the three substrates, L-hfOrn, L-Thr and L-hOrn as aminoacyl adenylates and subsequently transfer them onto the adjacent T domains to form the corresponding aminoacyl thioesters (Figure 3.18; 1). The α-carbons of the thioesters undergo epimerization catalysed by E domains of module 1 and 2. In the second step, the C domains of the module 2 and 3 catalyse sequential condensation of the thioesters of the three substrates, L-hfOrn, L-Thr and L-hOrn to form a tripeptide tethered to the T domain of module 3 (Figure 3.18; 2). In the third step, a second molecule of L-hfOrn is loaded onto T domain of module 1 and the α-carbon of the resulting thioester is epimerized by the module 1 E domain (Figure 3.18; 3). In the fourth step the C domain of either module 2 or module 3 catalyses the condensation of the module 3 attached tripeptide thioester with the module 1 attached hOrn thioester yielding tetrapeptide thioester attached to the T domain of module 3 (Figure 3.18; 4). In the final stage, the tetrapeptide thioester is released from the NRPS by hydrolysis catalysed by CchJ (Figure 3.18, 5), a protein encoded by one of the genes in the *cch* biosynthetic cluster. Usually the C-terminal TE domain is used by NRPS for the cleavage of the assembled peptide and the CchH was not found to contain one. However, the examination of other genes in the *cch* cluster identified an orthologue of an esterase, which catalyses hydrolytic cleavage of an enterobactin siderophore into smaller subunits in order to release iron from the ferric-enterobactin complex

(Brickman and McIntosh, 1992). Subsequent genetic knockouts of *cchJ* from *S. coelicolor* chromosome confirmed the role of CchJ as the free standing thioesterase that is involved in the release of the NRPS assembled tetrapeptide (Lautru et al., 2005).

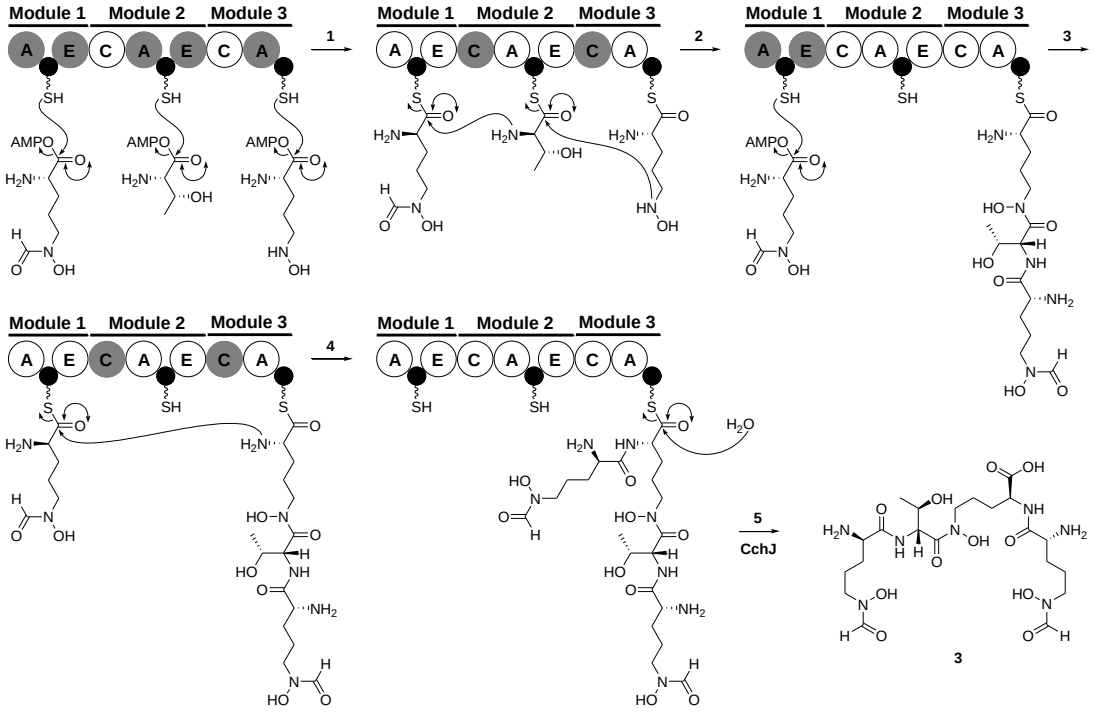


Figure 3.18 Proposed coelichelin biosynthetic pathway (Lautru et al., 2005). The stages 1-5 are described in the text. The domains active in each step are shaded gray. The black circles are the thiolation domains in each module of CchH.

In scabichelin biosynthesis the putative NRPS Scab85471 also lacks the TE domain. However, none of the other gene in the putative scabichelin biosynthetic gene cluster encodes the hydrolase protein.

The coelichelin biosynthetic model shows the iterative use of the A, T and E domain of module 1 and the C domain of either module 2 or 3 and some domains in the second assembly are skipped. Iterative module use and module skipping is the unique trait of coelichelin biosynthesis and the predicted assembly of scabichelin, a putative hydroxamate siderophore of *S. scabies*, does not follow that pattern.

Two genes from coelichelin biosynthetic cluster *cchA* and *cchB* encode the enzymes with sequence similarity to formyl transferase and NADH/FAD-dependent monoxygenases, respectively. CchB catalySe the hydroxylation of L-Orn to yield L-hOrn and CchA subsequently formylates it yielding L-fhOrn (Figure 3.19 A). These two non-proteinogenic amino acids are later incorporated into coelichelin. The role of both enzymes was experimentally confirmed in the biosynthesis of coelichelin (CchB) and rhodochelin (CchA; Pohlmann and Marahiel, 2008; Bosello et al., 2012).

Bioinformatic analysis of the functional domains of the NRPS encoded for by *scab85471* revealed it contains 17 enzymatic domains organised into 5 modules (Figure 3.19 B), hence the product of the *scab85471* gene would be a pentapeptide. On the basis of prediction of the substrate specificity of the adenylation domains in each module it was identified that modules 1, 2 and 3 of the NRPS likely select for a L-N5-acyl-N5-hydroxyornithine, a L-Ser and a L-N5-hydroxyornithine, respectively. The substrate specificity for the modules 4 and 5 could not be, however, unambiguously predicted.

Two genes from the putative scabichelin biosynthetic cluster, *scab85511* and *scab85521*, encode proteins homologous to CchA and CchB from *S. coelicolor* M145, respectively. *Scab85521* is proposed to encode the flavin-dependent monooxygenase that catalyse N5 hydroxylation of L-Orn to form L-N5-hydroxyornithine (L-hOrn). *Scab85511* encodes the putative formyl-tetrahydrofolate-dependent formyl transferase that catalyses N5-formylation of L-N5-hydroxyornithine to form L-N5-formyl-N5-hydroxyornithine (L-fhOrn; Figure 3.19 A). These reactions are predicted to be analogous to the reactions catalysed by

CchA and CchB, enzymes involved in the biosynthesis of non-proteogenic amino acid residues hOrn and fhOrn that are utilised in the biosynthesis of coelichelin (Pohlmann and Marahiel, 2008). The amino acids, which are the building blocks for the putative *S. scabies* product are loaded onto the thiolation (T) domains by the adenylation (A) domains of the NRPS encoded by *scab85471* (Figure 3.19 B). The methylation (MT) domains of the NRPS are proposed to catalyse the N-methylation of the aminoacyl thioesters attached to the T domains within module 1 and 3. The epimerization (E) domain within module 5 is proposed to catalyze α-carbon epimerization of the L-hOrn residue attached to the T domain of module 5 and the condensation (C) domains of modules 2 to 5 are proposed to catalyse peptide bond formation between the T domain-bound aminoacyl thioesters to yield pentapeptidyl thioester attached to the T domain of module 5 before it is finally released from the NRPS (Figure 3.19 B).

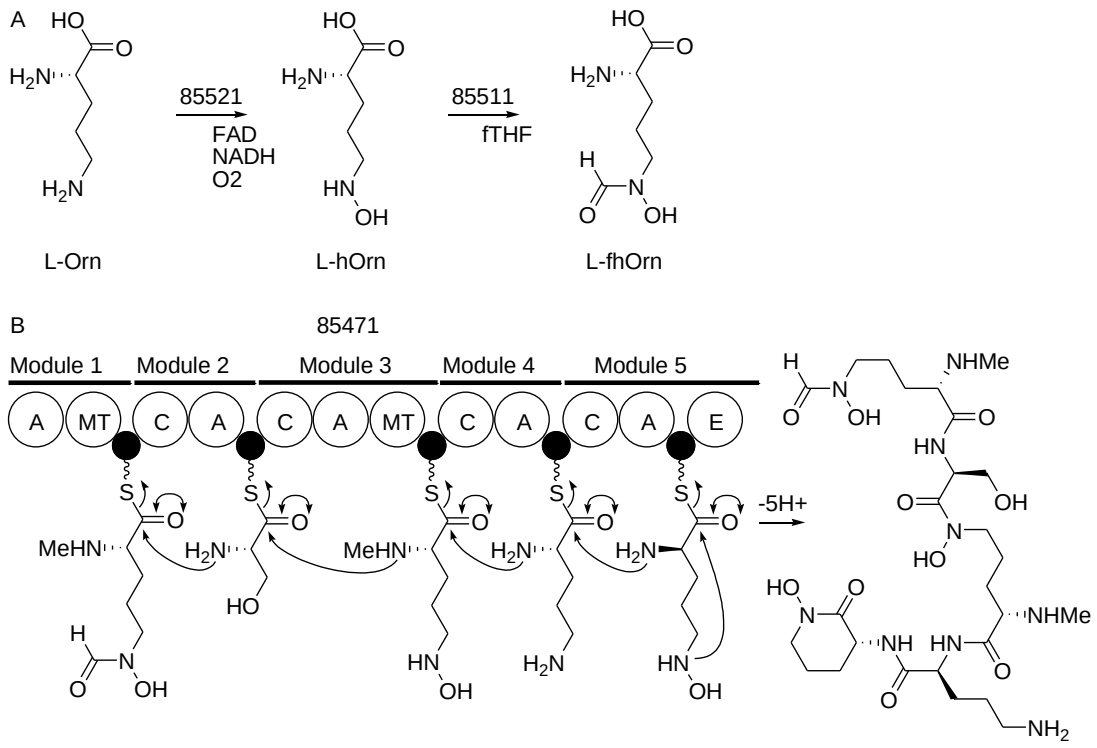


Figure 3.19 A-Proposed role of the enzymes encoded by *scab85521* and *scab85511* in the biosynthesis of the amino acids L-N5-hydroxyornithine (L-hOrn) and L-N5-formyl-N5-hydroxyornithine (L-fhOrn). B-Organisation of modules and domains of the NRPS encoded by *scab85471*. The residues attached to the thiolation (T) domains (black circles) of the five modules are the aminoacids proposed to be incorporated into scabichelin.

There are two hypothetical ways of scabichelin release from the NRPS. One of them is via intramolecular addition of the δ-hydroxylamino group of the D-hOrn residue to the carbonyl group of the thioester, and subsequent deprotonation and release of the resulting product (Figure 3.19 B). This mechanism for peptide chain release is analogous to that proposed for another putative siderophore called amychelin (Figure 3.21) from *Amycolatopsis* sp. AA4 (Seyedsayamdost et al., 2011). Both amychelin and scabichelin, contains D-cyclo-hOrn residue and both of the putative siderophore NRPSs lack the C-terminal thioesterase (TE) domain, which is responsible for chain release in nonribosomal peptide biosynthesis. In amychelin that reaction is proposed

to be catalysed by AcmB, a putative standalone α, β-hydrolase (Figure 3.20). Scabichelin biosynthetic gene cluster, however, does not encode a protein with sequence similarity to an α, β-hydrolase. Therefore, the reaction may be catalysed by α, β-hydrolase encoded elsewhere in the *S. scabies* genome. Some epimerization domains may also be involved in the peptide release. The epimerization (E) domain of module 5 of the scabichelin NRPSs may not only catalyse the epimerization of the module 5 bound L-hOrn residue, but it may also catalyse the subsesquent cleavage of the assembled peptide. *In vitro* experiments are required to elucidate the peptide release mechanisms.

Further work, in collaboration with Dr. Bignell (former member of professor R. Loria group at Cornell University) and co-workers, would be to examine the effect of scabichelin biosynthetic gene inactivation on plant pathogenicity and would involve the use of a plant assay to compare the necrotic effect of *S. scabies* 87.22 and W1000 strains grown in parallel on tuber and root vegetables such as potato and radish, respectively.

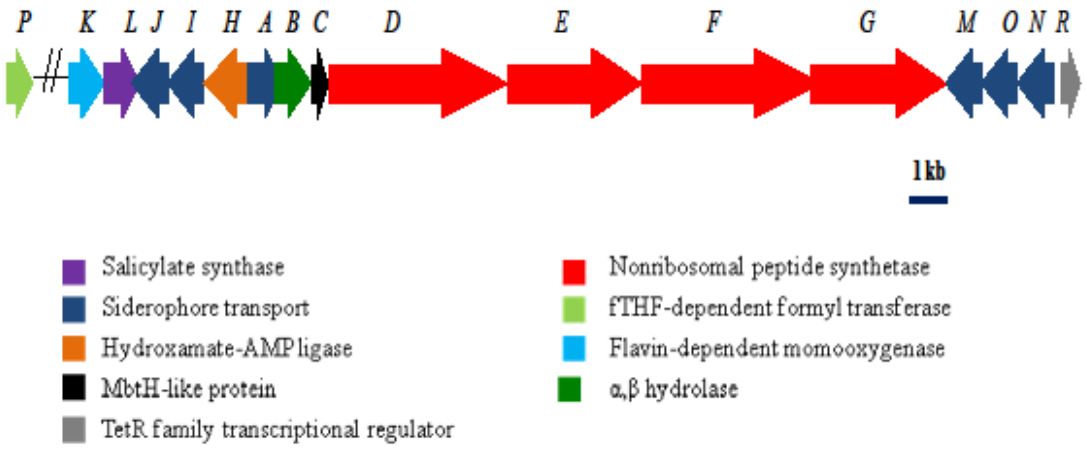


Figure 3.20 Amychelin biosynthetic gene cluster in *Amycolatopsis sp.* AA4 (Seyedsayamdost at al., 2011).

The amychelin biosynthetic gene cluster is 35.2 kb in size and consists of 16 genes (SSMG_02531 to SSMG_02546). It includes four NRPS genes (*amcDEFG*), amino acid tailoring genes (*amcKP*), chain initiation and termination genes (*amcBCHL*) and transport genes (*amcAIJMON*). AmcK is 61% identical to CchB in *S. coelicolor* a flavin-dependent monooxygenase and *amcP* is encoded outside of the *amc* gene cluster and it encodes a protein that is 74% identical to CchA the formyltransferase in coelichelin biosynthetic pathway. Thus, AmcK and AmcP are proposed to catalyse formation of the two nonproteinogenic precursor amino acids N-hydroxyornithine (hOrn) and N-hydroxy-N-formylornithine (hfOrn), respectively.

AmcL is proposed to encode salicylate synthase that catalyses the salicylate from chorismate in proposed amychelin biosynthesis (Figure 3.21; Seyedsayamdost et al., 2011). AmcH is proposed to encode hydroxybenzoyl-AMP ligase, which likely catalyses adenylation of salicylate and its subsequent loading onto the aryl carrier protein (ArCP) domain of AmcG. The substrate specificity of A domain in AmcG is predicted to be cysteine, though the serine is incorporated. The cyclisation domain (Cy) in AmcG is then proposed to catalyse oxazoline ring formation and the chain is then transferred to AmcF. The two A domains of AmcF have substrate specificity for serine and after the incorporation of second serine the E domain in AmcF catalyzes epimerization of second serine α-carbon. The chain is then passed on to AmcD, where it is condensed with another serine and with N-hydroxy-N-formylornithine (hfOrn) residue, a substrates predicted to be selected by the A domains of AmcD. The E domain in AmcD generates D form of hfOrn. The A domain of the final NRPS AmcE is predicted to select for glutaminic acid. After condensation of hfOrn and

epimerization of its α-carbon the peptide is proposed to be released from the AmcE by AmcB a putative α, β hydrolase encoded in *amc* cluster. AmcE does not contain the C-terminal TE domain and the amychelin release mechanisms resembles that of coelichelin, where free standing esterase performs that function (Seyedsayamdost et al., 2011).

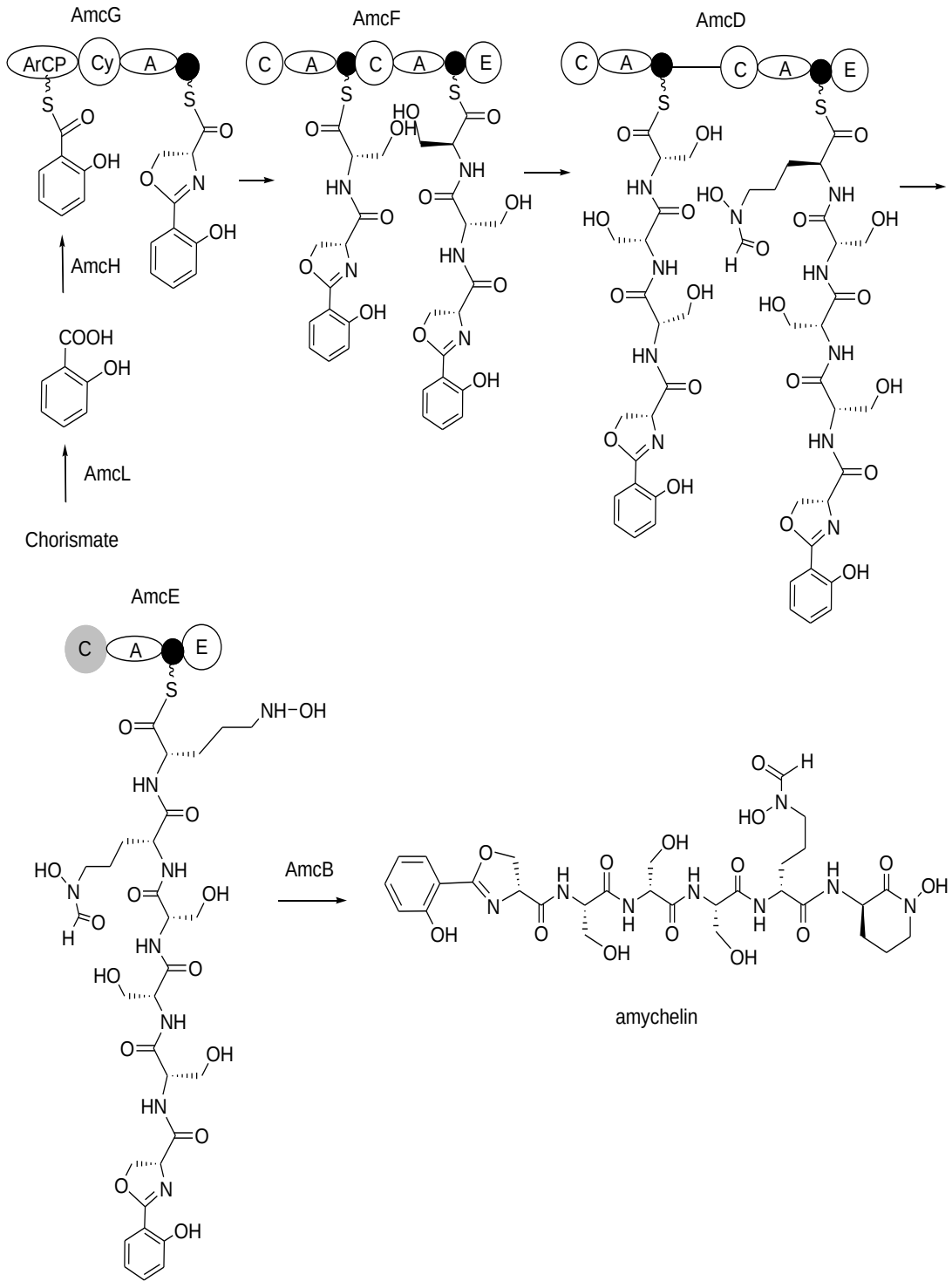


Figure 3.21 Proposed biosynthesis of amychelin (Seyedsayamdost at al., 2011).

4. Desferrioxamine production in *S. scabies* 87.22

4.1 The Aims of This Work

Desferrioxamines are a group of ferric iron binding molecules consisting of hydroxamic acid ligands (Challis and Kadi, 2009). Unlike pyochelin and scabichelin, which are synthesised by multidomain NRPS enzymes, desferrioxamines are synthesised via an NRPS-independent pathway (Challis and Kadi, 2009). Desferrioxamines are produced by a broad range of bacteria including many *Streptomyces* species (Bickel et al., 1960) and analysis of the *S. scabies* genome unveiled a putative desferrioxamine biosynthetic gene cluster.

The aim of this work was to establish if the putative desferrioxamine biosynthetic gene cluster in *S. scabies* directs the biosynthesis of desferrioxamine and to investigate the role of the *desC* gene in desferrioxamine biosynthesis. This was achieved by creating a *desC* gene deletion in *S. scabies* using the PCR targeting technology and examining the production of desferrioxamine in the *desC* mutant strain. Subsequent chemical complementation of the *desC* mutant by the introduction of a precursor from the desferrioxamine biosynthetic pathway into *desC* mutant liquid culture was analysed to establish if desferrioxamine production could be restored. The *desC* mutant in *S. scabies* was also genetically complemented by *in trans* expression of the *desC* gene from the integrative plasmid pOSV556.

4.2 Putative Desferrioxamine Biosynthetic Gene Cluster in *S. scabies*

Analysis of the *S. scabies* 87.22 genome sequence identified a cryptic biosynthetic gene cluster, which includes the protein coding sequences Scab57921-Scab57981 (Figure 4.1). These proteins are homologs of proteins encoded by genes in desferrioxamne biosynthetic gene cluster of *S. coelicolor* A (3)2 (Table 4.1).

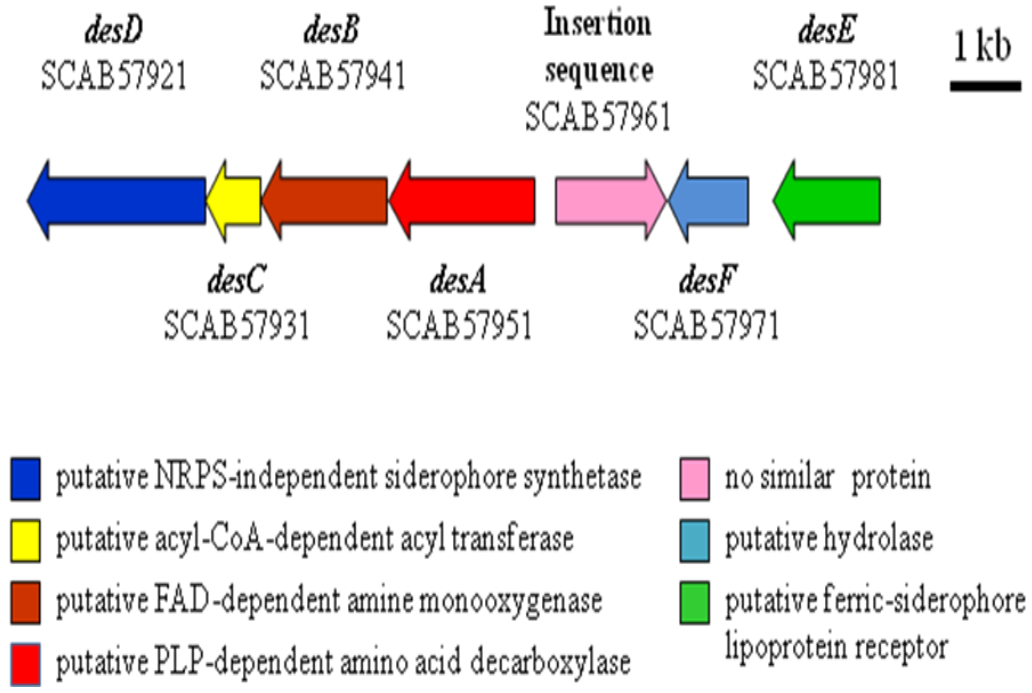


Figure 4.1 Putative desferrioxamine biosynthetic gene cluster in *S. scabies* 87.22.

Table 4.1 Putative proteins functions encoded by the desferrioxamine biosynthetic gene cluster in *S. scabies* 87.22.

***S. scabies* gene**	**desferrioxamine biosynthesis protein (Des)**	**Sequence similarity with *S. coelicolor* A3(2) *des* cluster (%)**
scab57921	DesD	hypothetical protein SCO2785 (81%)
scab57931	DesC	aceytltranferase SCO2784 (79%)
scab57941	DesB	monooxygenase SCO2783 (80%)
scab57951	DesA	pyridoxal-dependent decarboxylase SCO2782 (86%)
scab57961	IS630 family insertion sequence	no similar protein
scab57971	DesE	hypothetical protein SCO2781 (75%)
scab57981	DesF	secreted protein SCO2780 (73%)

The *desC* gene, which is of interest, encodes for an enzyme responsible for catalysing the acylation of N-hydroxycadaverine to form N-hydroxy-N-succinylcadaverine, a precursor in desferrioxamines biosynthesis in *Streptomyces coelicolor* (Kadi amd Challis, unpublished data; Figure 4.2).

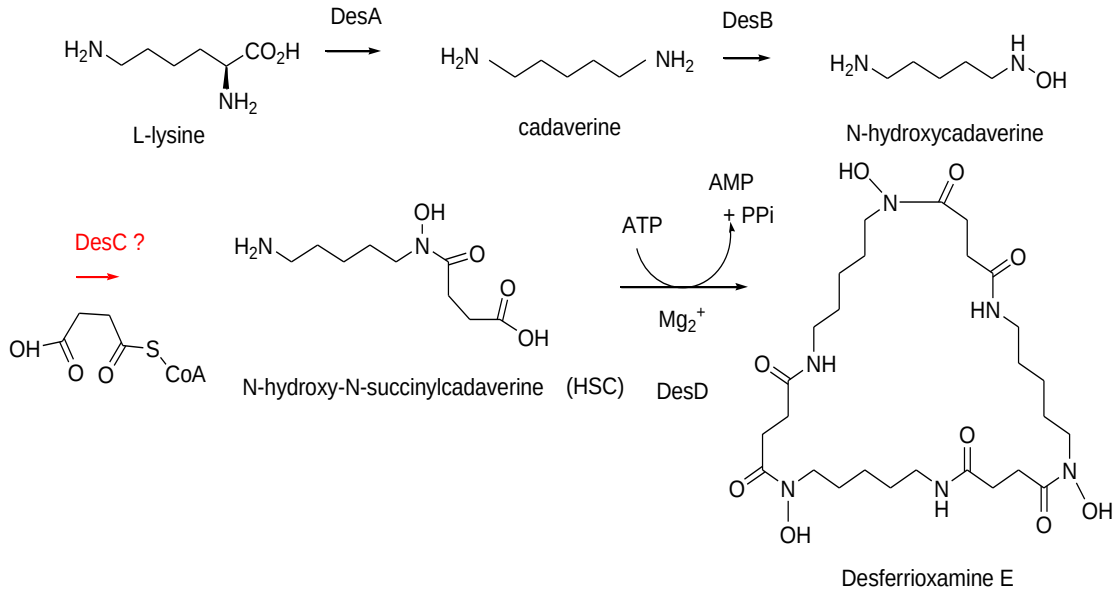


Figure 4.2 Proposed desferrioxamine biosynthetic pathway in *S. coelicolor* (Kadi et al., 2007).

4.3 Construction of *des*C Mutant in *S. scabies* 87.22

The *desC* mutant was previously made with *S. scabies* ATCC 49173 (Loria group, unpublished data). However, before continuing with the *desC* complementation experiment, it has been decided that the *desC* mutant should be made again in *S. scabies* 87.22, which is the wild type strain used by our collaborators at Cornell University. The reason for making *desC* mutant in *S. scabies* 87.22 was to investigate function of that gene and desferrioxamine production in this strain. Specifically after the *desC* mutant of *S. scabies* 87.22 has been made, the complementation of this mutant was attempted.

The *desC* mutant in *S. scabies* 87.22 was created using the same strategy as that used to create the scabichelin mutant i.e. PCR targeting and subsequent conjugation to the wild type *S. scabies* 87.22 strain. The cosmid containing the whole *des* cluster with the *oriT*-apramycin resistance cassette in place of the *desC* gene, which is only 555 bases in length, was used for the construction of the mutant (the sequence of *des* cluster and the cosmid was obtained from our collaborators). This cosmid was introduced into *E. coli* ET12567 by electroporation and then transferred by conjugation from *E. coli* ET12567 to *S. scabies* 87.22. The kanamycin sensitive, apramycin resistant *Streptomyces* colonies were then selected and their genomic DNA was isolated and analysed by PCR. The PCR (Figure 4.3) was performed using 18 nt long test primers (Table 7.6) designed within 100 bp region upstream and downstream of *desC* gene in *S. scabies* 87.22 genome.

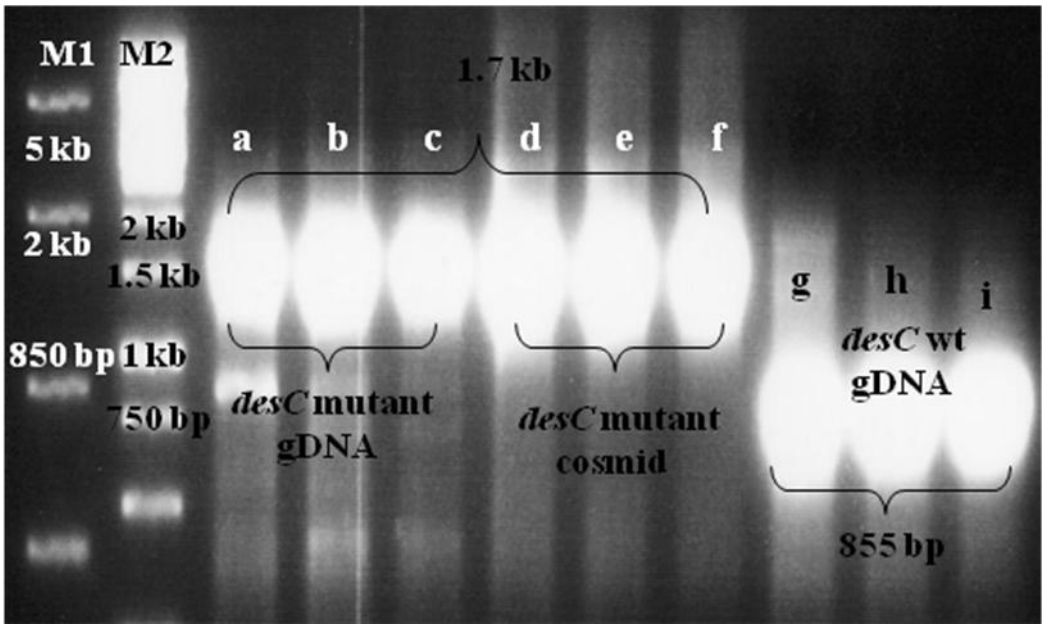


Figure 4.3 PCR analysis of genomic DNA from the putative *desC* mutant colony. Lanes: M1- Middle Range DNA molecular size marker; M2- 1kb DNA molecular size marker, a, b, c- *desC* mutant genomic DNA; Tm= 57, 61, 63,respectively; d, e, f- *desC* cosmid with apramycin cassette; g, h, i- wild type *S. scabies* 87.22 genomic DNA.

A PCR product of approximately 1.7 kb in size was obtained in the mutant strain. This size for the mutant corresponds to the apramycin cassette of 1.4 kb plus the primers priming about 150 base pairs downstream and upstream of the cassette. This PCR product was excised and purified from the gel and sent for sequencing analysis. The sequencing results confirmed that the PCR product contains the apramycin resistance gene. The size of the PCR product for the wild type *desC* gene was 855 bp.

Southern blot hybridization analysis was carried out to determine if the resistance cassette was inserted in the correct position in the genome. The DIG-labeled *des* cluster cosmid was used as a probe. Genomic DNA and the cosmid containing the

des cluster were digested using *Sac*I. The probe was hybridized to the digested DNA and the different pattern of bands for the wild type and mutant were present on X-ray film (Figure 4.4).

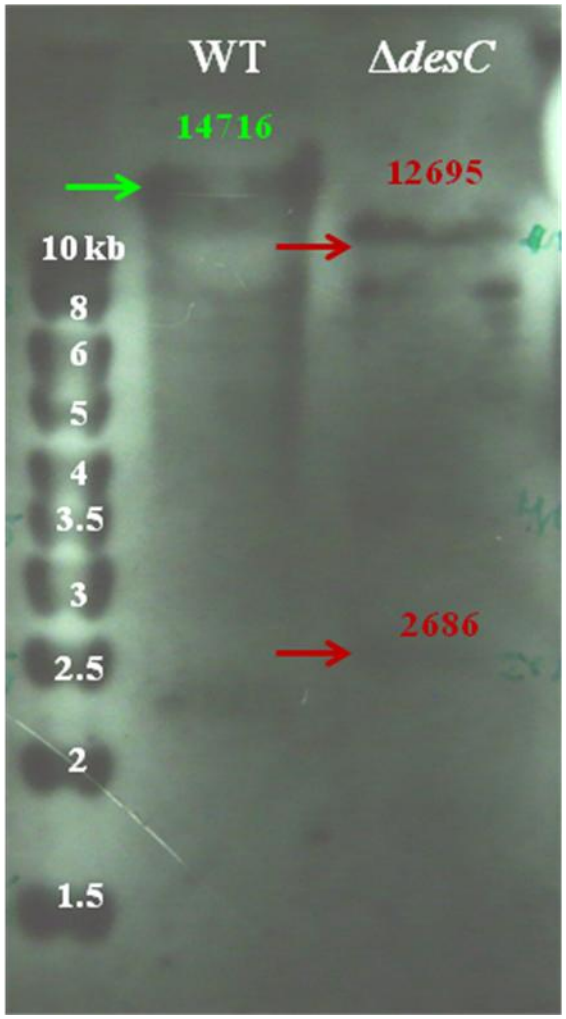


Figure 4.4 Southern blot hybridization using labeled *des*::*oriT-apr* cosmid as a probe confirming the nature of *S.scabies desC* mutant. Band 14716 bp in green is only present in wild type strain. Bands 1412 bp and 7575 bp coloured in red are characteristic for the mutant.

4.3.1 LC-MS Analysis of *desC* Mutant

The metabolic profile of the *desC* mutant was analysed by LC-MS. The *desC* mutant strain was grown in 50 ml Minimal Liquid Medium (SMM) in parallel with 50 ml SMM culture of *S. scabies* 87.22 wild type at 30° C, 180 rpm for 5 days. The culture extracts were then analysed via LC-MS for the presence of ferrioxamine B and E. Analysis of the extracted ion chromatograms for the *m/z* of 614 (ferrioxamine B; Figure 4.5 A) and 654 (ferrioxamine E; Figure 4.5 B) indicated that these compounds are present in cultures of the wild type strain. However, compounds with the same *m/z* and retention time could not be seen in the *desC* mutant (Figure 4.5 A, B).

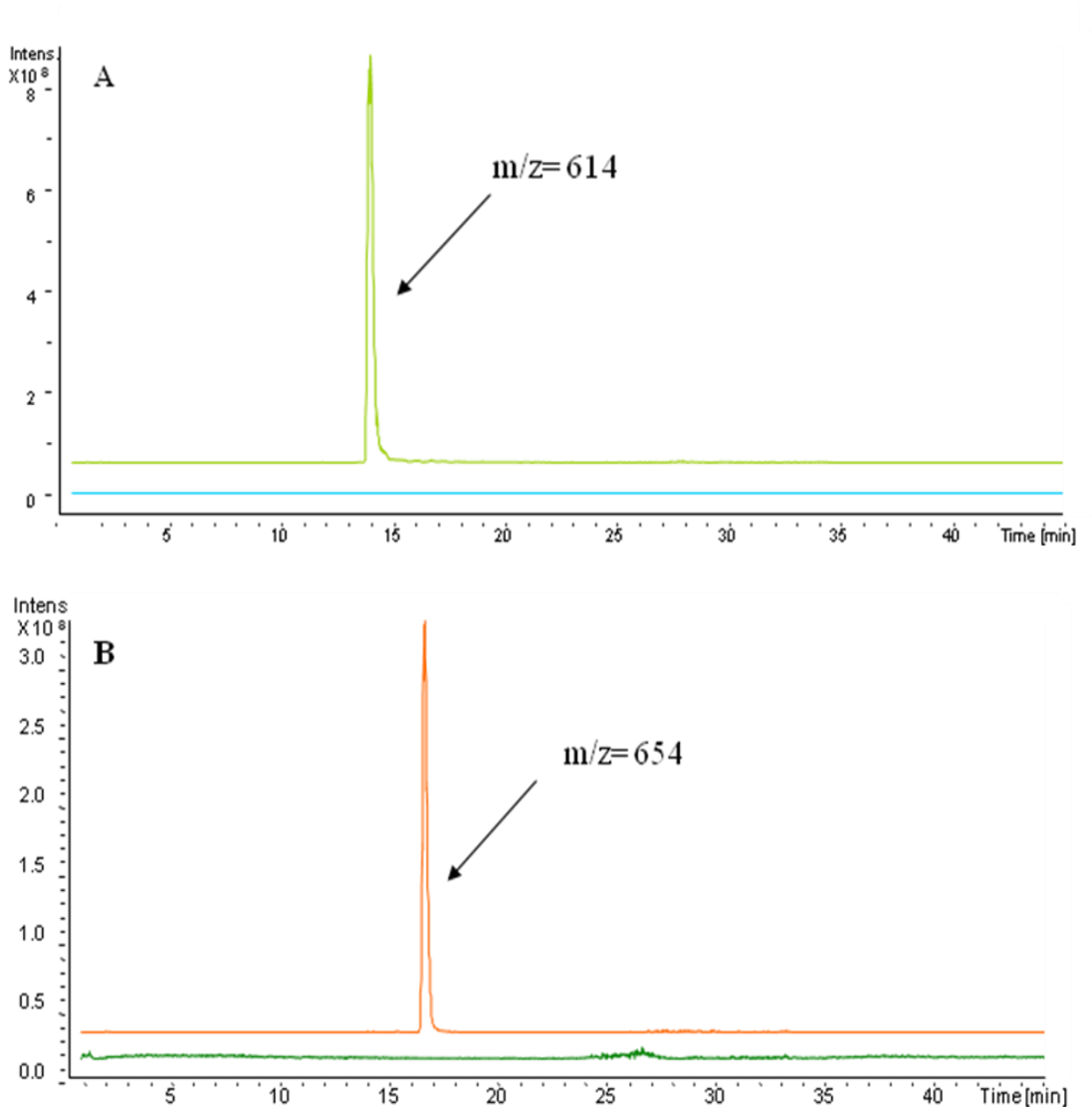


Figure 4.5 A – EIC (*m/z* = 614) for ferrioxamine B from LC-MS analyses of acidified organic extracts from the *S. scabies* wild type (green) and *S. scabies desC* mutant (blue). B – EIC (*m/z* = 654) for ferrioxamine E from LC-MS analyses of acidified organic extracts from the *S. scabies* wild type (red) and *S. scabies desC* mutant (green).

4.4 Chemical Complementation of *desC* Mutant

The aim of this work was chemical complementation of the *S. scabies des*C mutant to illustrate that reintroduction of the N-hydroxy-N-succinylcadaverine (HSC) precursor (Figure 4.6) into the desferrioxamine biosynthetic pathway would restore the production of desferrioxamines in this mutant. This was carried out by setting up three tubes with 5 ml of SMM medium each. Two were inoculated with the *desC* mutant spores, and *S. scabies* 87.22 wild type strain was used to inoculate the third. The HSC precursor was added to one of the cultures in which the *desC* mutant was growing. The other culture with *desC* mutant and the one with wild type *S. scabies* remained as the controls with no HSC added to either one. Following 72 hours incubation the culture supernatants were analysed by LC-MS. The analysis indicated restoration of desferrioxamine B production in the culture of *desC* mutant chemically complemented with HSC (Figure 4.7). However, desferrioxamine E (*m/z*=654) was not detected in the culture of *desC* mutant chemically complemented with HSC. Desferrioxamine was not detected in the culture supernatant of *desC* mutant with no HSC added (Figure 4.7).

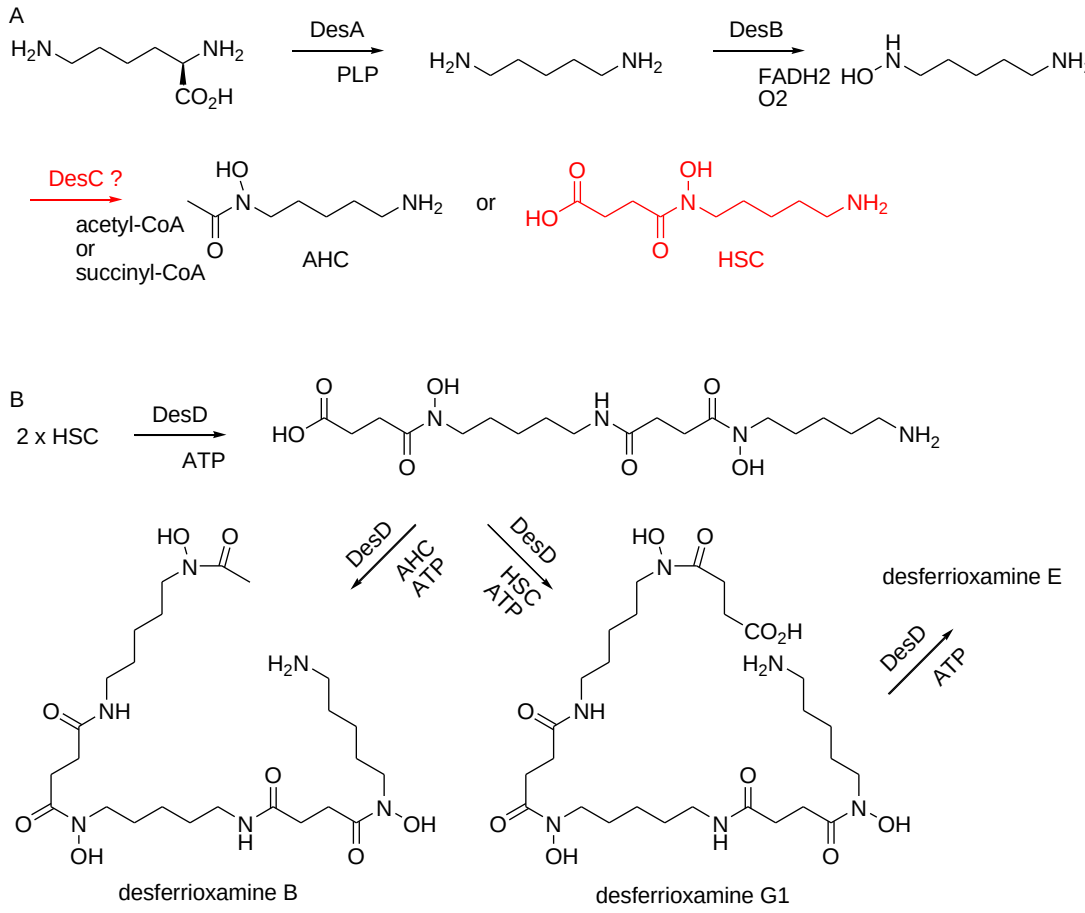


Figure 4.6 Pathway for the biosynthesis of desferrioxamines B, G1 and E (Kadi et al., 2007).A-reactions catalysed by DesA, DesB enzymes and reaction proposed to be catalysed by DesC enzyme. B-Reaction catalysed by DesD enzyme. yielding desferrioxamine B, G1 and E.

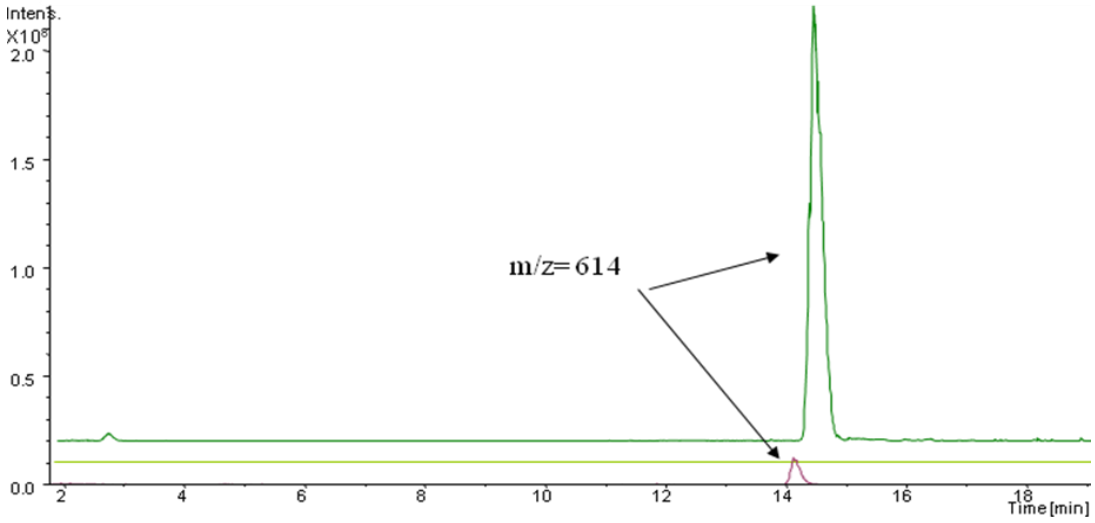


Figure 4.7 EIC (*m/z* = 614) for ferrioxamine B from LC-MS analyses of acidified organic extracts from the *S. scabies* wild type (top trace), *S. scabies desC* mutant (middle trace) and Δ *desC* mutant + HSC (bottom trace).

4.5 Genetic Complementation of *des*C Mutant

Genetic complementation of the *des*C mutant by in *trans* expression of the deleted gene from the integrative plasmid pOSV556 (Figure 4.8), in addition to the chemical complementation, confirmed that reintroduction of the *desC* gene restores the production of desferrioxamines in *S. scabies*. *desC* gene complementation in *S. scabies* was carried out by cloning the *desC* gene into pOSV556 plasmid, which is under the control of the constitutive *ermE** promoter. The plasmid contains ampicillin resistance gene (for selection in *E.coli*) and the hygromycin resistance gene (for selection in *Streptomyces*). It also contains *oriT*, allowing for the conjugal transfer of the plasmid from *E. coli* to *Streptomyces*. Additionally, an integrase encoded within pOSV556 allows its integration into the *Streptomyces* chromosome.

It occurs through site-specific recombination between the *attP* site within the vector and the *attB* site within the chromosome (Raynal et al., 2002).

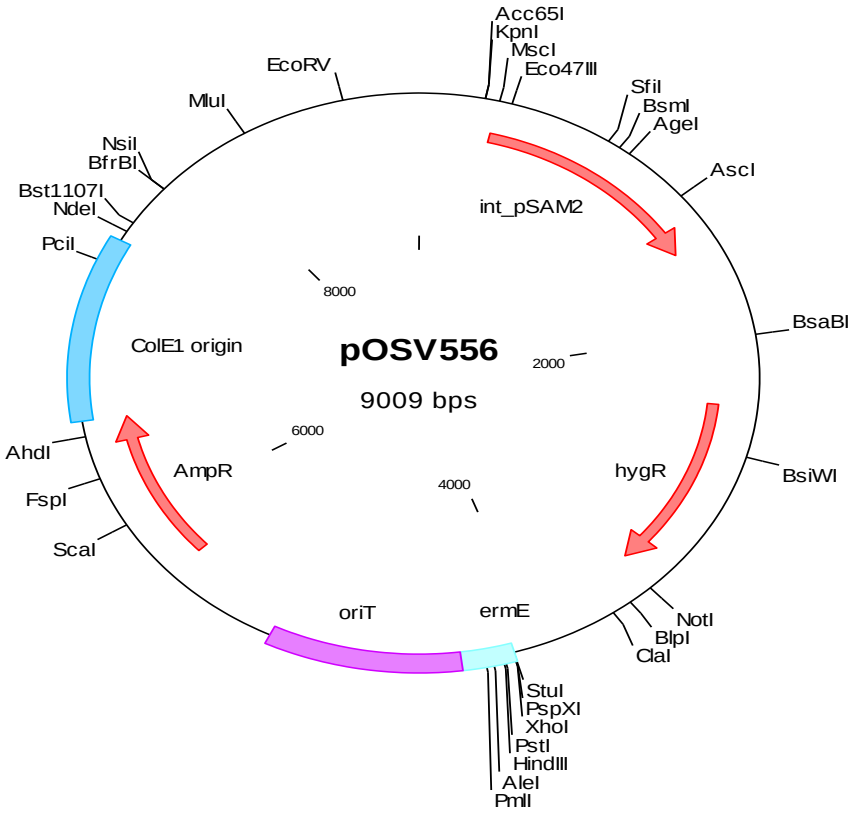


Figure 4.8 Map of pOSV556 plasmid.

To clone the *desC* gene into pOSV556 PCR cloning primers were designed (section 7.1.5). The forward primer included an artificial conserved ribosome binding site (RBS; 5'-AGGAGG-3') followed by the 6 bp upstream of the start codon and the first 7 bp of the coding sequence (Strohl, 1992). Reverse primers were designed to anneal approximately 100 bp downstream of the stop codon of the gene. The forward primer contained a *Hin*dIII restriction site and the reverse primer contained an *Xho*I site. The amplified *desC* gene was then cloned in the multiple cloning site of

pOSV556. The resultant PCR product was digested with *Hin*dIII and *Xho*I restriction enzymes and subsequently ligated with the similarly digested pOSV556. Electrocompetent *E. coli* DH5α were then transformed with the ligation mixture and colonies were selected for on LB agar containing ampicillin. Correct insertion of the PCR product into the vector was then confirmed by restriction digest and by PCR using the cloning primers (Figure 4.9). The size of a product was 699 bp. The *E. coli* ET12567/pUZ8002 were transformed with the correctly identified plasmid clone, which was subsequently transferred by conjugation from *E. coli* ET12567/pUZ8002 to *S. scabies* strain and a single hygromycin resistant exconjugant was chosen for further analyses. DNA was isolated and analysed by PCR for the insertion of the *desC* gene copy from the pOSV556 by the cloning primers (Table 7.6; Figure 4.9).

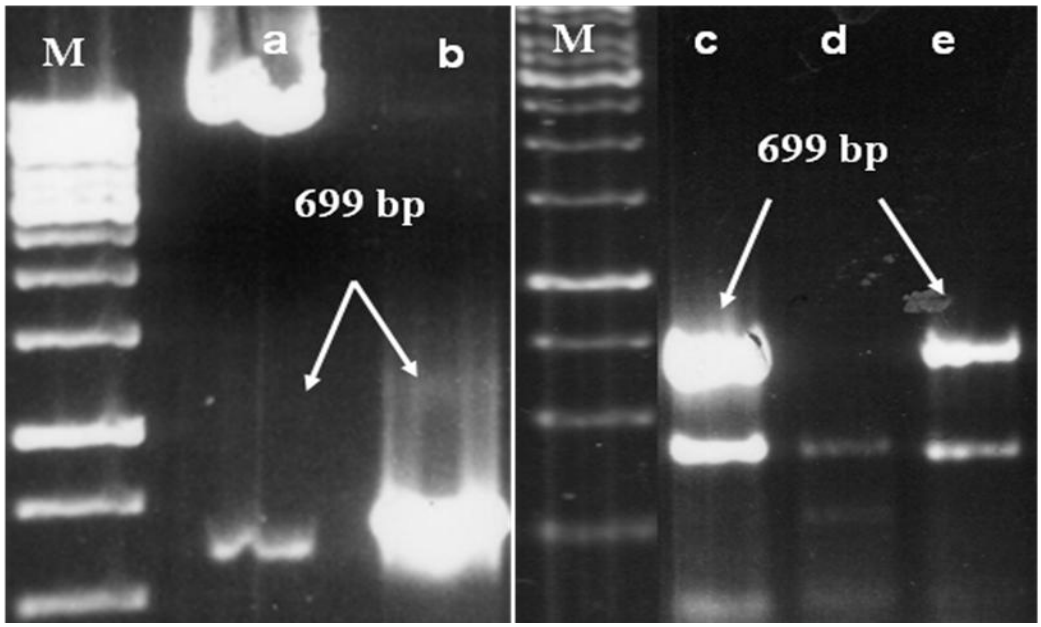


Figure 4.9 Left- a- an *XhoI-HindIII* restriction digest and b- PCR analysis with *desC* gene cloning primers of pOSV556::*desC* plasmid. Right-PCR analysis with *desC* gene cloning primers of: c- *S. scabies* wild type chromosomal DNA, d- *S. scabies* Δ*desC* mutant DNA and e- *S. scabies* Δ*desC::*pOSV556::*desC* complemented mutant. M- 1 kb molecular DNA ladder.

4.5.1 LC-MS Analysis of Complemented Δ *des*C Mutant

The LC-MS analysis (Figure 4.10) indicated restoration of desferrioxamine B production in the culture of the *desC* mutant complemented with pOSV556*::desC,* however, this is at a lower level when in comparison to the *S. scabies* 87.22 wild type strain. Desferrioxamine B (*m/z*=654) was not detected in the culture of the *desC* mutant complemented with pOSV556*::desC*. Desferrioxamine was not detected in the culture supernatant of *desC* mutant (Figure 4.10).

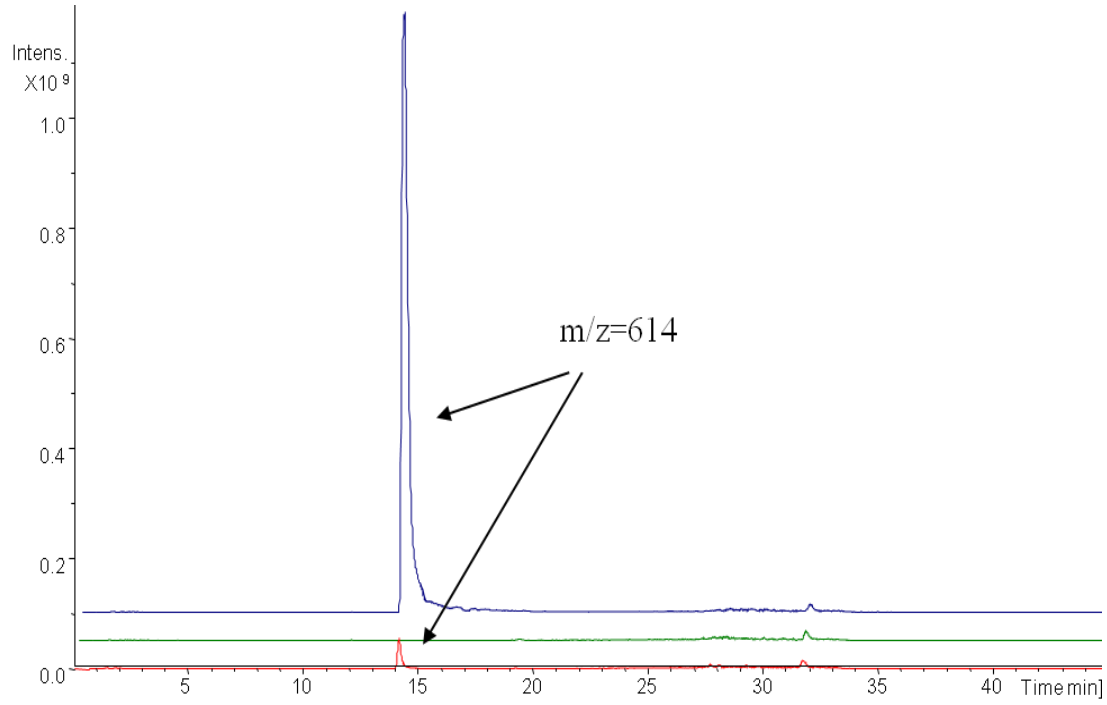


Figure 4.10 EICs (*m/z* = 614) for ferrioxamine B from LC-MS analyses of acidified organic extracts from the *S. scabies* wild type (top trace), *S. scabies desC* mutant (middle trace) and Δ *desC* ::pOSV556/63021::apr (bottom trace).

4.6 Discussion

The PCR-targeting method was used to create a deletion of the *desC* gene within the *des* cluster of *S. scabies* (Gust et al., 2002). The *desC* gene in *S. scabies* genome was replaced with an apramycin resistance cassette containing *oriT*. The constructed mutant was analysed via PCR and Southern blot hybridisation methods.

LC-MS analysis of the *S. scabies* wild type strain and the *S. scabies* Δ*desC* mutant strain liquid cultures was undertaken. Before analysis, the supernatants of the wild type and the mutant strain were incubated with ferric iron to obtain the ferric complexes with desferrioxamines, which were then examined via LC-MS. The results indicated that wild type *S. scabies* produced desferrioxamines, whereas the Δ*desC* mutant *S. scabies* strain did not, indicating that disruption of the *desC* gene in *S. scabies* abolished desferrioxamine production.

Chemical complementation was carried out to further analyse the function of *desC* gene within the *des* cluster in *S. scabies*. The DesC enzyme is thought to catalyse the acylation of N-hydroxycadaverine to form precursors in desferrioxamine E and B biosynthesis in *S. coelicolor* (N. Kadi and G. L. Challis unpublished data; Figure 4.6).

Chemical complementation using HSC restored desferrioxamine B production in the *S. scabies desC* mutant, although to a lower level than in a wild type *S. scabies* strain. The lower level of desferrioxamine B production in *S. scabies desC* chemically complemented strain could have been due to the amount of HSC precursor added to *desC* mutant liquid culture. The concentration of HSC in *desC*

mutant culture was only 1µmol/ml. The reason for the small concentration of HSC was that it was not readily available during the experiment. Had more HSC been available the chemical complementation of *S. scabies desC* mutant could have been repeated several times to validate the LC-MS analysis results and optimise the concentration needed.

The introduction of *desC* copy from pOSV556 vector into the Δ*desC* mutant *S. scabies* genome was confirmed by PCR and LC-MS analyses. The LC-MS of complemented Δ*desC* mutant strain indicated restoration of desferrioxamines biosynthesis.

Investigation into the role of DesC was carried out in *S. coelicolor* (N. Kadi and G. L. Challis, unpublished data). The *desC* gene encodes an enzyme similar to acyl-CoA-dependent acyl transferases (Kadi et al., 2007), and catalyses the N-acylation of *N*-hydroxycadaverine (Figure 4.6; N. Kadi and G. L. Challis, unpublished data). In those organisms that also produce desferrioxamine B, DesC has relaxed substrate specificity, allowing acylation of N-hydroxycadaverine with acetyl-CoA as well as succinyl-CoA, giving N-hydroxy-N-acetylcadaverine (AHC) and N-hydroxy-N-succinylcadaverine (HSC), respectively (Figure 4.6; Kadi et al., 2007). To prove directly that DesC catalyses these reactions the DesC enzyme purified from *S. coelicolor* was incubated with succinyl coenzyme A and with N-hydroxycadaverine substrate in one reaction and with acetyl coenzyme A and with N-hydroxycadaverine in another reaction (Kadi and Challis, unpublished data). LC-MS analysis of the reaction mixture with succinyl coenzyme A and another with acetyl coenzyme A showed respectively the presence of the peaks, which had the same *m/z* and retention

times as the authentic standards of the products HSC and AHC (Kadi and Challis, data not shown).

In conclusion, *desC* gene in *S. scabies* is involved in desferrioxamine biosynthesis as its deletion resulted in the abolition of desferrioxamine production. Chemical complementation of the *desC* mutant with HSC restored to some level desferrioxamine biosynthesis (Figure 4.10) suggesting that the DesC protein in *S. scabies* catalyses the acylation of N-hydroxycadaverine to yield HSC, which is utilised in subsequent reaction catalysed by DesD to yield desferrioxamine B.

Results and Discussion II: Investigation of Tetronate-like Biosynthetic Gene Cluster in *S. scabies* 87.22.

5. Investigation of Tetronate-like Biosynthetic Gene Cluster in *S. scabies* 87.22.

5.1 Cryptic Polyketide Tetronate Biosynthetic Gene Cluster in *S. scabies* 87.22

S. scabies 87.22 was found to encode a cryptic tetronate biosynthetic gene cluster (Figure 5.1). It contains genes homologous to the genes in the RK-682 biosynthetic cluster, such as a type I polyketide synthase, an Acyl-CoA synthetase, an Acyl carrier protein, a FabH-like protein and a FkbH-like protein. The metabolic product of this gene cluster has not yet been identified. It is hypothesised to be a novel compound with a structure similar to that of RK-682 and agglomerins, and with a potential role in plant pathogenicity.

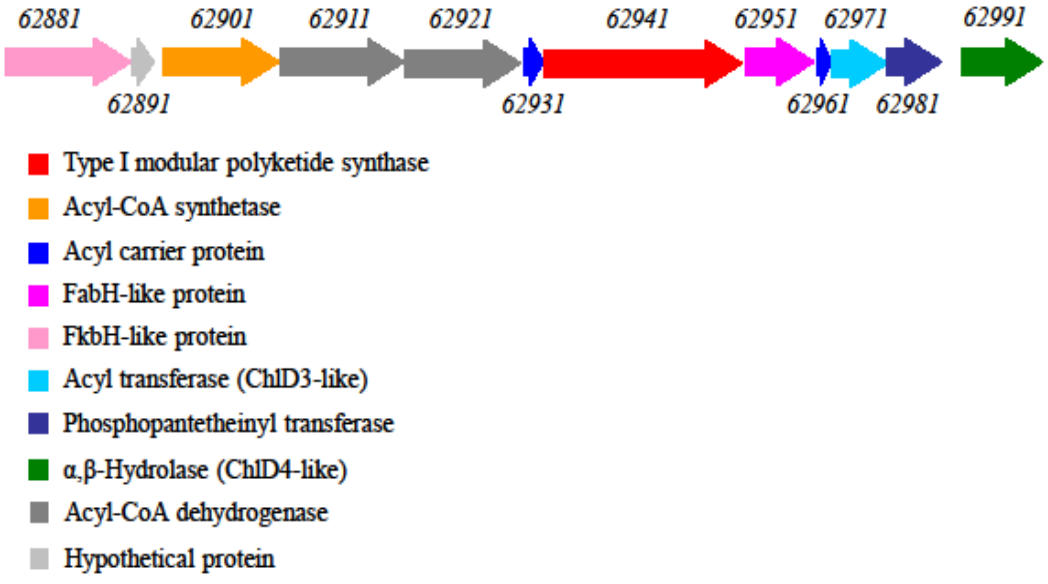


Figure 5.1 Cryptic tetronate polyketide biosynthetic gene cluster Figure in *S. scabies* (Challis, unpublished).

5.2 Transcriptional Analysis of Cryptic Tetronate Biosynthetic Gene Cluster in *S. scabies*

The aim of this work was to determine if the genes from the putative tetronate-like cluster in *S. scabies* are expressed under laboratory conditions. Currently the metabolic products of this cluster have not been identified. To check the expression of the putative tetronate-like cluster in *S. scabies* reverse transcriptase PCR (RT-PCR) was carried out on the total RNA isolated from wild type *S. scabies*.

PCR primers approximately 20 bp in length were designed to amplify a 500-600 bp regions of several genes found within the putative tetronate biosynthetic gene cluster, such as *scab62881* (encoding a FkbH-like protein), *scab62951* (encoding a FabH-like protein), *scab62971* (encoding a ChlD3-like protein), *scab62991* (encoding a ChlD4-like protein) and also from the *hrdB* gene (an essential gene encoding sigma-like transcription factor) as a positive control. The total RNA was isolated from *S. scabies* mycelium grown for 3, 5, 7 and 14 days. The RNA was then analysed using agarose gel electrophoresis (Figure 5.2) and treated with DNAse to remove all traces of DNA from the RNA samples. RT-PCR was carried out on a purified RNA with the *hrdB, scab62881, scab62951, scab62971* and *scab62991* primers (Section 7.1.5) to examine the transcription of the genes (Figure 5.3).

The first step in RT-PCR involves the use of reverse transcriptase polymerase and a primer to anneal and extend the targeted mRNA sequence. The mRNA is transcribed to a complementary strand of DNA (cDNA), which is subsequently amplified with primers and DNA polymerase in a standard PCR. The PCR products are analysed using agarose gel electrophoresis. If a band of the desired molecular weight is

present, then the mRNA was present in the sample, consequently the associated gene is expressed.

The figure 5.3 B shows the molecular band present for the *hrdB* control PCR was the correct size. Analysis of the RT-PCR for the tetronate genes *scab62881* and *scab62951* indicates that the tetronate-like cluster is expressed as bands of the correct size were present (Figure 5.3 C, D). However, the PCR results for *scab62991* and *scab62971* are inconclusive (data not shown). The RT-PCR product obtained for *scab62991* was not the correct size and in the control reaction, PCR on genomic DNA, multiple bands could be seen, thus indicating that the primers for this gene are not specific enough and did not anneal to the correct RNA sequence.

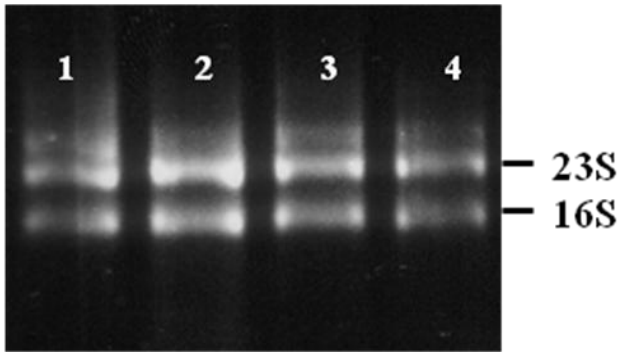


Figure 5.2 The 1% agarose gel of total RNA extracted from wild type *S. scabies*. Lanes 1-4: 1- RNA extracted from 3 days old mycelium; 2- RNA extracted from 5 days old mycelium; 3- RNA extracted from 7 days old mycelium; 4- RNA extracted from 14 days old mycelium.23S and 16S are two bands of total RNA.

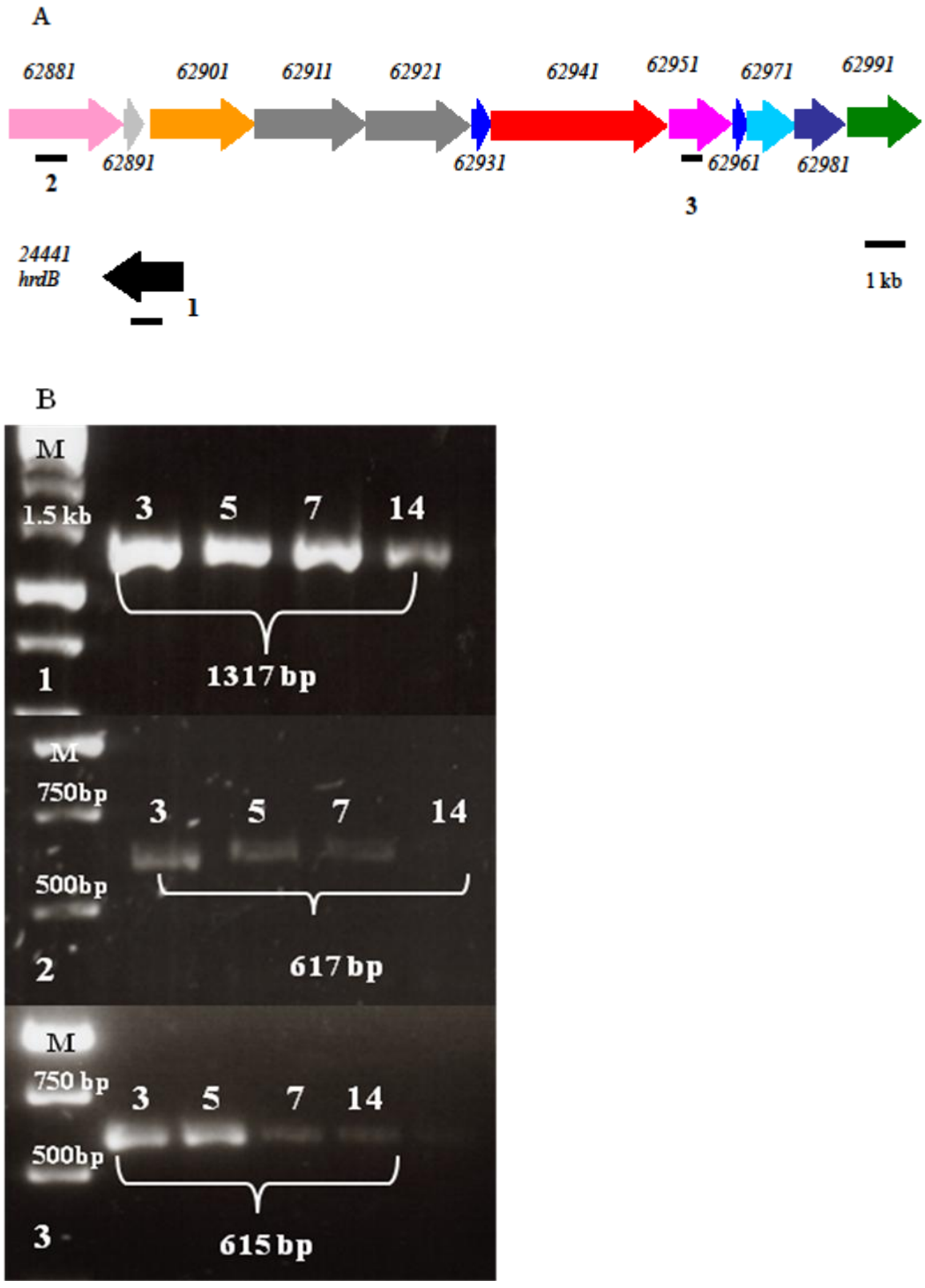


Figure 5.3 RT-PCR analysis of *S. scabies* wild type RNA. A- the putative tetronate gene cluster in *S. scabies* 87.22; B- RT-PCR: 1- with *hrdB* primers (positive control); 2- with *scab62881* primers; 3- with *scab62951* primers. 3, 5, 7, 14- the *S. scabies* cells were grown for 3, 5, 7 and 14 days before the RNA extraction.

5.3 Mutagenesis of *scab63021* to Determine the Effect on Expression of Putative Tetronate-like Gene Cluster in *S. scabies*.

5.3.1 SARP Family

Scab63021 is proposed to encode for a transcriptional activator containing some sequence homology to the SARP (Streptomyces Antibiotic Regulatory Proteins) family of genes. SARP genes are pathway-specific regulatory genes that are associated with individual biosynthetic gene clusters and they were the first pathway-specific regulatory proteins to be discovered (Wietzorrek and Bibb, 1997). Proteins of the SARP family are transcriptional activators. Examples include RedD and ActII-orf4, which regulate the production of undecylprodigiosin and actinorhodin, respectively, in *Streptomyces coelicolor* (Fernandez-Moreno et al., 1991; Narva and Feitelson, 1990). It is proposed that SARPs activate transcription by binding to a set of sequences of heptameric repeats located in the proximity of the cognate promoters of biosynthetic clusters, which are controlled by SARPs (Garg and Parry, 2010). When the SARPs are bound to the promoters it is proposed that the mechanism of transcription initiation involves binding of RNA polymerase to the appropriate sites (Tanaka et al., 2007).

5.3.2 Mutagenesis of *scab63021* Gene

The *scab63021* gene is situated approximately 1500 bp downstream of the *S. scabies* tetronate-like biosynthetic gene cluster. It has been proposed that it is involved in the regulation of the tetronate-like biosynthetic gene cluster acting as a transcriptional

activator. To prove this hypothesis the gene was disrupted in *S. scabies* using PCR targeting protocol (section 7.6.5).

In this method the cosmid containing the target gene is usually used. However, the cosmid containing *scab63021* was not available. Instead, primers were designed to encompass the start and stop codons of the target gene and to amplify it by PCR from *S. scabies* 87.22 genomic DNA. The amplified *scab63021* gene would be then cloned into plasmid vector and that construct could subsequently be used to transform *E.coli* BW25113/pIJ790 strain. From then on the further steps in the construction of the mutant would be carried as in a standard PCR Redirect protocol.

According to the above the cloning primers were designed to amplify *scab63021* gene via PCR. These were designed in the same way as the primers used to clone the *S. scabies desC* gene into pOSV556 plasmid (section 4.6), however, the PCR with primers designed to amplify the *scab63021* gene did not yield a product with the correct molecular weight. For that reason the *scab63021* gene was not cloned into plasmid vector. As the alternative, the *scab63021* gene was synthesised by a gene synthesis company (Genewiz). The whole *scab63021* gene was synthesised and cloned into pUC57 plasmid by Genewiz (Figure 5.4).

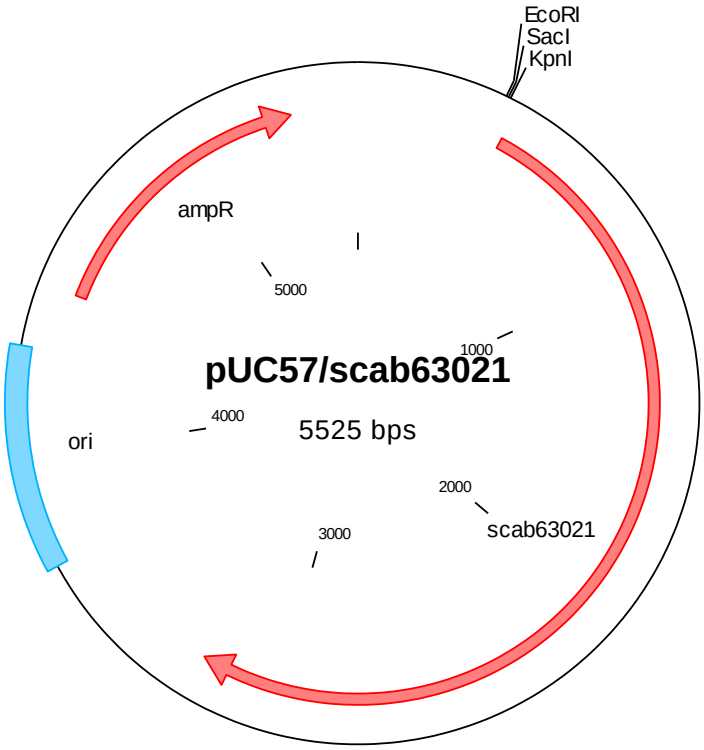


Figure 5.4 Map of pUC57/*scab63021*. Legend: *oriT*-origin of replication, *ampR*-ampicillin resistance gene.

E.coli BW25113/pIJ790 was electrotransformed with the obtained pUC57/ *scab63021* construct. A 1 kb internal fragment of *scab63021* was chosen to be replaced. To disrupt the *scab63021* gene (size 2781 bp) in *S. scabies*, specific primers (Section 7.1.5) were designed and used to amplify the *oriT-apramycin* cassette obtained from the pIJ773 plasmid. The priming sequences (39 nt) used were taken from 880 bp after the start codon and about 740bp before the stop codon of the targeted gene i.e., within the *scab63021*. Each primer had at the 5' end 39 nt homologous to the *S. scabies* sequence adjacent to the targeted 1 kb sequence inside of the *scab63021* and at the 3' end 19 or 20 nt matching the right or left end of the disruption cassette (section 3.1). The *E. coli* BW25113/pIJ790 bacteria containing

the pUC57 *scab63021* construct were electrotransformed with the PCR amplified *OriT*-apramycin cassette resulting in a double cross-over recombination between the ends of the amplified resistance cassette and the homologous regions within *scab63021*. To determine if the desired gene replacement had occurred correctly, the pUC plasmid was isolated and analysed by PCR using test primers, which were designed to anneal ~100 bp upstream/downstream of the region in the *scab63021* gene to be replaced. The expected band for mutagenised plasmid was 1770 bp, and for wild type plasmid 1440 bp. As it is illustrated in Figure 5.5 the PCR for the wild type plasmid did not work very well but it showed a band with a size of about 1700 bp for the mutated plasmid had been amplified as expected. This band was purified from the gel and sent for sequencing, which indicated the presence of the apramycin resistance gene. The replacement of the *scab63021* gene fragment in the plasmid was also confirmed via restriction enzymes digestion patterns of the mutated plasmid with the wild type plasmid (examples of digestion with *Sac*I, which cuts twice within the *oriT-apr* cassette generating a specific 752 bp band indicative of the presence of the cassette, and *Bam*HI enzymes are shown in Figure 5.6).

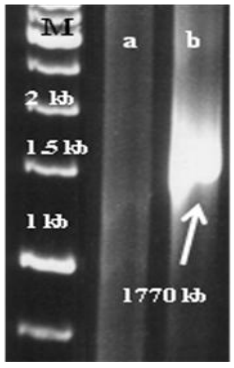


Figure 5.5 PCR analysis of genetically-engineered pUC57/*scab63021* plasmid with test primers. Lanes: M-b. M- 1 kb molecular weight marker; a- PCR band for unmutagenised pUC57/*scab63021*; b- PCR for pUC57/*scab63021::oriT-apra*.

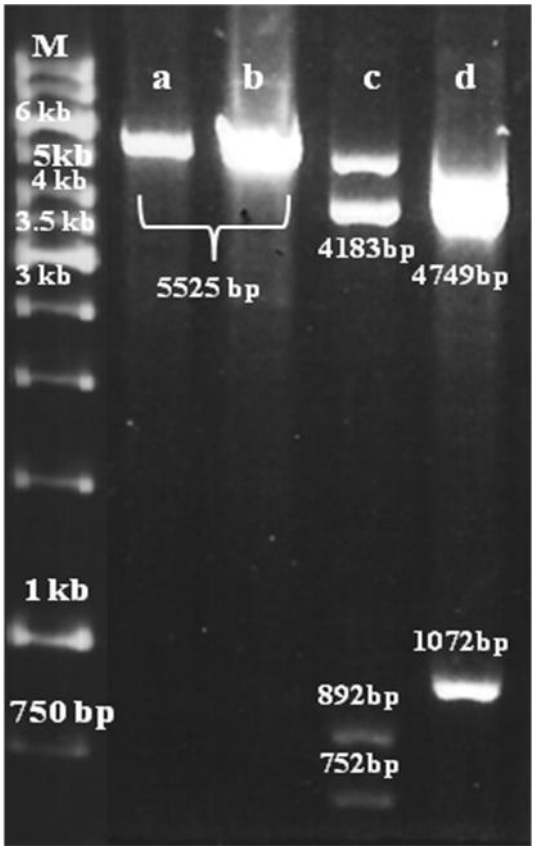


Figure 5.6 Agarose gel electrophoresis analysis of restriction enzymes digest of genetically-engineered pUC57/*scab63021* plasmid used to disrupt the *scab63021* in the *S. scabies* genome within putative tetronate-like cluster. Lanes M-d: M- 1kb molecular weight marker; a- undigested pUC57/*scab63021::oriT-apra*; b- digestion with *Eco*RI*;* c- digestion with *Bam*HI d- digestion with *Sac*I.

The correctly modified plasmids were introduced into *S. scabies* by intergenic conjugation from *E. coli* ET12567/pUZ8002. Apramycin resistant, *Streptomyces* exconjugants were analysed by PCR to confirm that a double cross-over gene replacement had taken place (Figure 5.7). The first PCR using genomic DNA from the *scab63021* mutant colonies and with the primers annealing inside of the *oriT-apramycin* cassette, confirmed the presence of the *oriT*-apramycin cassette in place of the *scab63021* internal fragment in the genome of the mutant. The expected size of the PCR product was ~400 bp, which was then compared to the PCR of the

mutagenised plasmid pUC/*scab63021::apr*, which also gave band of ~400 bp (Figure 5.7 A). This indicate that introduction of *oriT-apramycin* cassette into the genome of putative *scab63021* mutant was successful. They do not prove, however, that a double cross-over between *scab63021* and cassette had occured. The putative *scab63021* mutant genomic DNA was further analysed with the test primers (Section 7.1.5) designed to anneal ~100 bp upstream and downstream of the replaced internal fragment within *scab63021*. The expected band size was ~1.5 kb for the wild type *S. scabies* genomic DNA, and 1.8 kb for *scab63021* mutant. The PCR results showed no band for wild type, however, the expected band was present for the mutant strain (only mutant band is displayed-Figure 5.7 B).

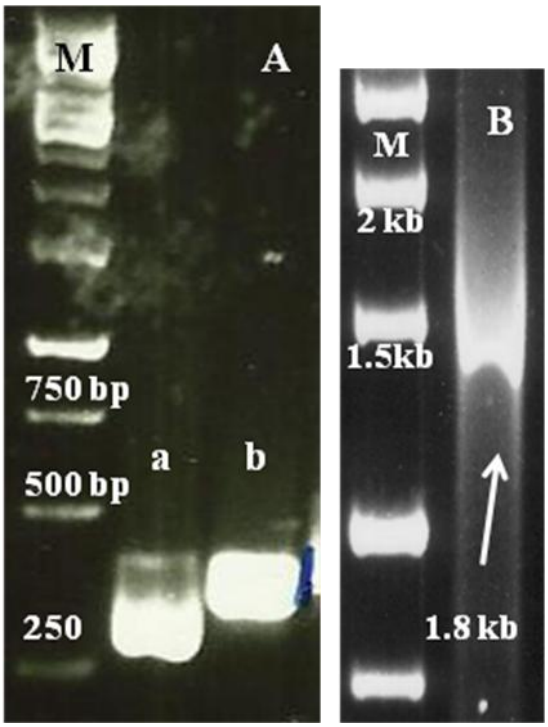


Figure 5.7 A- PCR analysis of genetically-engineered pUC57/*scab63021* plasmid and genomic DNA of *scab63021* mutant with *oriT- apramycin* casette internal primers; M-1 kb molecular weight marker, a –PCR of genomic DNA of *Δscab63021* mutant; b- pUC57/*scab63021*::apra plasmid with apramycin internal primers. B-PCR analysis of putative *scab63021* mutant genomic DNA with test primers.

5.4 RT-PCR on Wild Type *S. scabies* and the Δ*scab63021* Mutant RNA

Although the mutation of the *scab63021* gene had not been proved conclusively, analysis of the wild type *S. scabies* and the putative *scab63021* mutant genomes has been carried out to allow for comparisons in the patterns of expression of the genes in the putative tetronate-like gene cluster.

The RT-PCR was carried on wild type and mutant *S. scabies* purified RNA using the same method previously described (section 5.3). The RNA was extracted from mycelia 3, 5, 7 and 14 days old. In this experiment the expression was analysed only for two genes from *S. scabies* tetronate-like cluster; *scab62881* (encoding a FkbH-like protein), and *scab62951* (encoding a FabH-like protein) as the RT-PCR in wild type *S. scabies* yielded specific bands only for these two genes within the cluster. The RT-PCR with the *scab62881* primers yielded specific as well as unspecific bands. The right size for the specific band of RT-PCR product on wild type RNA was 617 bp for the *scab62881* primers. The ~20 bp PCR primers designed for the previous analysis to amplify regions of *scab62881* and *scab62951* approximately 600 bp in length were used in this experiment. The *hrdB* gene RT-PCR was again used as a positive control.

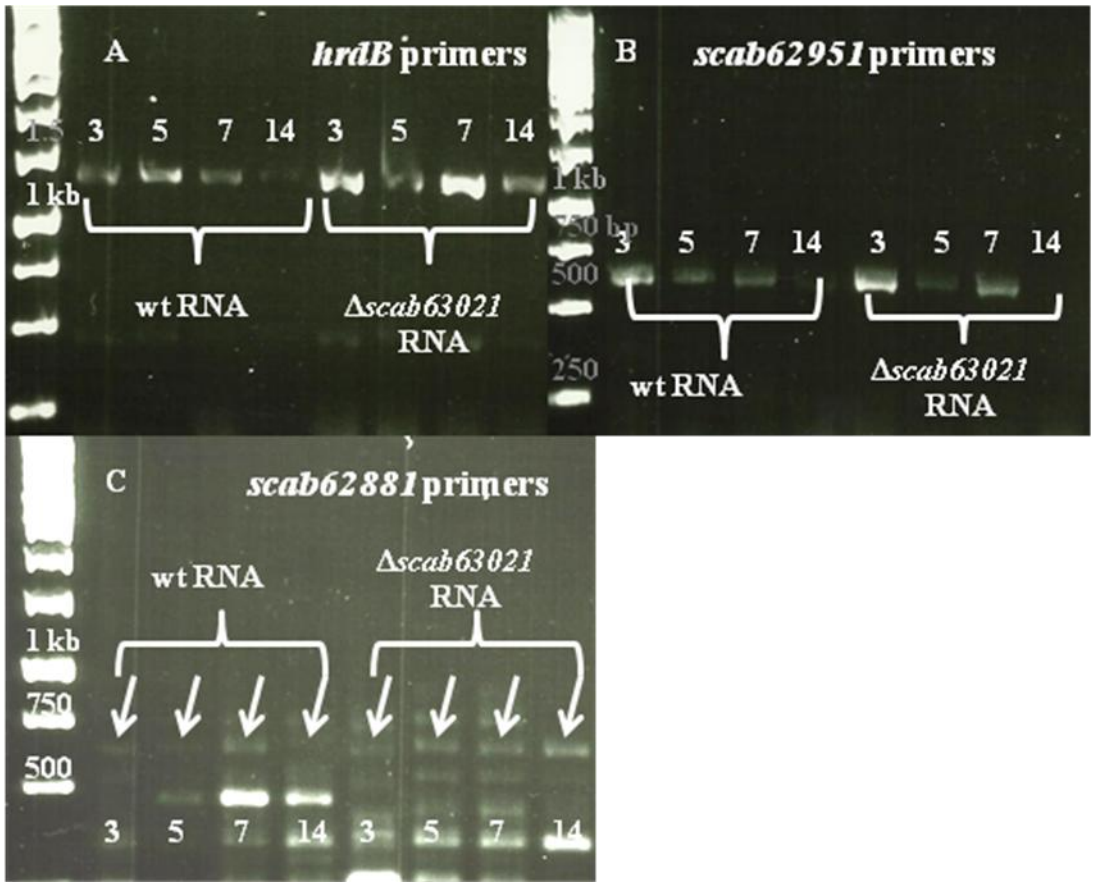


Figure 5.8 RT-PCR analysis of *S. scabies* wild type and Δ*scab63021* RNA, A: RT-PCR control, 1317bp product size (*hrdB* primers); B: RT-PCR with *scab62951* primers, 615 bp product size; C: RT-PCR with *scab62881* primers, 617 bp product size.

5.5 Discussion

The PCR analysis and sequencing of the Δ*scab63021* mutant genomic DNA has confirmed the presence of the resistance cassette in the mutant genome. However, this did not enable us to conclude that the cassette was in the correct position within the genome. For that Southern blot hybridisation was carried out. However, findings were inconclusive due to a high amount of background on the X-ray film.

The expression pattern does not illustrate any differences between the wild type strain and the Δ*scab63021* mutant indicating that *scab63021* does not have any effect on the expression of genes within the putative tetronate biosynthetic gene cluster.

An alternative method could have been employed to construct the Δ*scab63021* mutant. The *scab63021* gene could have been obtained within a cosmid, which could subsequently be used as a template for PCR. The cosmid would be more useful as the insert is bigger and the PCR primers could be designed to anneal outside of the *scab63021* gene and the sufficient flanking sequences on both sides of the target sequence would ensure that the homologous recombination between the investigated gene and the resistance cassette PCR product ends would occur.

Further work would be to carry out the comparative analysis of the *S. scabies* wild type strain and Δ*scab63021* mutant culture supernatants via LC-MS to find the putative tetronate product. If there would be no differences between the metabolic profiles of the two strains, the Δ*scab63021* mutant could be constructed again by another method. Alternatively, the deletion of the *scab62941*, encoding for a putative PKS within the *S. scabies* tetronate-like gene cluster could be obtained and the metabolic profile of the wild type strain and the mutant could be compared again via LC-MS in search of the cryptic tetronate metabolite.

6. Summary, Conclusions and Future Work

6.1 Results and Discussion I: Investigation of Siderophore Biosynthetic Gene Clusters in *S. scabies*.

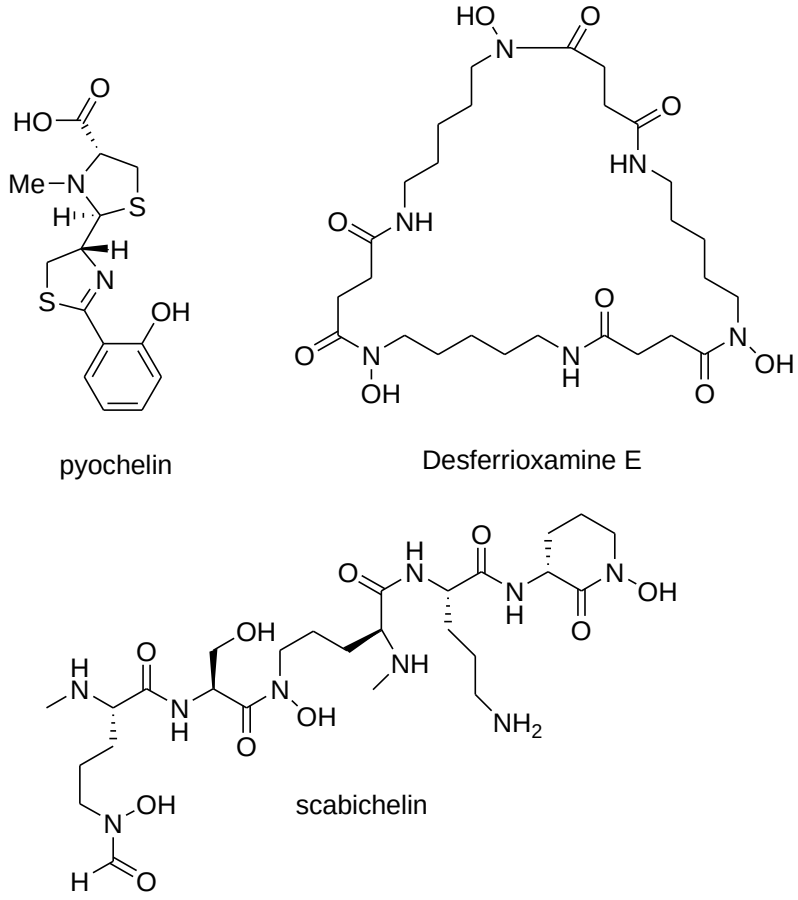


Figure 6.1 Structures of investigated siderophores.

6.1.1 Investigation of Pyochelin Biosynthetic Gene Cluster in *S. scabies* 87.22.

Putative functions for most of the twelve genes within the putative pyochelin biosynthetic gene cluster (PBGC) in *S. scabies* 87.22 have been proposed (Seipke *et al*., 2011) and a biosynthetic pathway for pyochelin has been suggested on the basis of bioinformatics and experimental studies. However, the regulation of pyochelin

production was not fully understood. *Scab1401* was found to encode a pathway specific regulator, which was predicted to repress transcription of PBGC in iron rich conditions. To determine the function of *scab1401*, the gene was replaced with the *oriT*-apramycin cassette creating *Δscab1401* mutant by our collaborators (Loria's group; Seipke et al., 2011). The mutant was then analysed via LC-MS for pyochelin production in iron deficient conditions. To prove that pyochelin is a product of *S. scabies* PBGC the putative NRPS encoding gene *scab1471* in *Δscab1401* mutant was replaced with *oriT*-hygromycin resistance cassette creating a *Δscab1401-Δscab1471* double mutant. The metabolic profile of the mutant was analysed via LC-MS alongside of *Δscab1401* mutant and the *S. scabies* wild type strain.

LC-MS analyses of the Δ*scab1401* mutant supernatant unveiled the presence of a compound with the same molecular mass as pyochelin. The stereochemistry of the compound produced by *S. scabies* was also investigated. Comparison of the Δ*scab1401* mutant culture supernatant by LC-MS using chiral stationary phase with the authentic standards of pyochelin and ent-pyochelin proved that *S. scabies* produce pyochelin, not enantio-pyochelin. The LC-MS analysis of culture supernatants of a Δ*scab1401*-Δ*scab1471* double mutant indicated no pyochelin production. These data indicate that pyochelin is produced by PBGC in *S. scabies* and Scab1401 represses transcription of pyochelin in iron-deficient conditions. The compound with the same mass as pyochelin was not consistently detected in the supernatants of wild type *S. scabies* when grown in iron-deficient conditions. It was postulated to be that it is due to that the pyochelin gene cluster in the wild type strain was not induced enough under the used growth conditions.

Plant assays were used (by Prof R. Loria's group at Cornell University; Seipke et al., 2011) to analyse the importance of pyochelin in plant pathogenicity. Incubation of Δ*scab1471* mutant (*scab1471* encodes one of the NRPS enzymes from the pyochelin biosynthetic pathway) and of *S. scabies* wild type, both on potato tissue and radish seedlings, indicated that pyochelin is not required for pathogenicity of the potato tissue and radish. One conclusion that can be drawn from these results is that iron may have not been essential for *S. scabies* growth in the assays. The genome of *S. scabies* possesses several more biosynthetic gene clusters encoding other siderophores, such as putative genes for the biosynthesis of desferroxamines (*scab57921–scab57951*). To determine which iron-chelating metabolites produced by *S. scabies* are required for plant pathogenicity, analysis of desferrioxamine mutant and other siderophores mutants in *S. scabies* is needed as well as, in planta transcriptional analysis of the pyochelin, desferrioxamine and other putative siderophore biosynthetic gene clusters.

6.1.2 Investigation of Scabichelin Biosynthetic Gene Cluster in *S. scabies* 87.22.

Mining of the *S. scabies* 87.22 genome has identified a gene cluster, which was predicted to encode the nonribosomal peptide synthetase multienzymes responsible for the biosynthesis of a novel iron-chelating metabolite named scabichelin. Bioinformatic analysis of a cryptic NRPS biosynthetic gene cluster, which include the protein coding sequences from *scab85431* to *scab85521* are homologous to proteins encoded for by genes within the coelichelin biosynthetic gene cluster from

S. coelicolor M145. Analysis of the culture supernatants of *S. antibioticus*, used by our collaborators, via HPLC and subsequent analysis by NMR indicated that these strains produce a tris-hydroxamate siderophore, which was named scabichelin. The aim was to establish whether the product of *S. antibioticus* is also the product of a cryptic NRPS encoded by the cryptic biosynthetic gene cluster in *S. scabies* 87.22.

Bioinformatic analysis of the functional domains of the NRPS encoded for by *scab85471* revealed it contains 17 enzymatic domains organized into 5 modules. Therefore, it was hypothesised that product of the *scab85471* gene would be a pentapeptide,

PCR analysis and sequencing confirmed the presence of apramycin cassette in *scab85471::oriT-apr* (W1000 mutant). The correct construction of the mutant was also confirmed by Southern blot hybridisation analysis.

Comparative LC-MS/MS analyses of the W1000 mutant, *S. scabies* 87.22 wild type and scabichelin purified from *S. antibioticus* indicated that scabichelin is the metabolic product of the *S. scabies* 87.22 cryptic gene cluster. LC-MS analyses of culture supernatants of the 87.22 and W1000 strains grown in an iron-deficient medium, to which ferric iron was added before analysis, indicated the presence of a ferric complex that is produced by *S. scabies* 87.22, but not the W1000 mutant. LC-MS retention time and MS/MS fragmentation pattern of ferri-scabichelin are the same for the *S. scabies* 87.22 compound and the authentic standard.

Bioinformatic analysis of the *scab85431-scab85521* gene cluster and LC-MS analysis of *S. scabies* 87.22 scabichelin indicated its potential function as an iron chelator. However, the essentiality of that metabolite for growth in iron-deficient

conditions has not yet been investigated. These findings may be helpful to establish the function of that siderophore in *S. scabies* and its potential role in plant pathogenicity. The further work with scabichelin would be conducted in collaboration with Dr.Bignell's research group (former member of Prof. Loria's group) and would involve the use of a plant assay to compare the necrotic effect of *S. scabies* 87.22 and the W1000 strains grown in parallel on tuber and root vegetables such as potato and radish, respectively.

6.1.3 Desferrioxamine Production in *S. scabies*

The PCR-targeting method was used to create a deletion of the *desC* gene within the *des* cluster of *S. scabies* (Gust *et al*., 2002). The *desC* gene in the *S. scabies* genome was replaced with an apramycin resistance cassette containing *oriT*. The constructed mutant was analysed via PCR and Southern blot hybridisation methods.

LC-MS analysis of the *S. scabies* wild type strain and the *S. scabies* Δ*desC* mutant strain liquid cultures was undertaken. Before analysis, the supernatants of the wild type and the mutant strain were incubated with ferric iron to obtain the ferric complexes with desferrioxamines, which were then examined via LC-MS.

Chemical complementation was carried out to further analyse the function of the *desC* gene within the *des* cluster in *S. scabies*.

The introduction of the *desC* copy from the pOSV556 vector into the Δ*desC* mutant *S. scabies* genome was confirmed by PCR and LC-MS analyses. The results indicated that the wild type *S. scabies* produced desferrioxamines, whereas the Δ*desC* mutant *S. scabies* strain did not, indicating that disruption of the *desC* gene in

S. scabies abolished desferrioxamine production. Chemical complementation using desferrioxamine pathway precursor N-hydroxy-N-succinylcadaverine (HSC) restored desferrioxamine B production in the *S. scabies desC* mutant, although to a lower level than in a wild type *S. scabies* strain. The lower level of desferrioxamine B production in *S. scabies desC* chemically complemented strain could have been due to the amount of HSC precursor added to Δ*desC* mutant liquid culture. The concentration of HSC in Δ*desC* mutant culture was only 1 µmol/ml. The reason for the small concentration of HSC was that it was not readily available during the experiment.

The LC-MS of genetically complemented Δ*desC* mutant strain indicated restoration of desferrioxamine biosynthesis.

The further work with *desC* mutant would be conducted in collaboration with Prof. Rose Loria's research group and would involve use of a plant assay to compare the necrotic effect of the *S. scabies* 87.22 wild type and Δ*desC* mutant grown in parallel on root vegetables.

6.2 Results and Discussion II: Investigation of Tetronate-like Biosynthetic Gene Cluster in *S. scabies* 87.22.

6.2.1 Investigation of Tetronate-like Biosynthetic Gene Cluster in S. scabies 87.22.

The *S. scabies* 87.22 genome was found to encode cryptic tetronate biosynthetic gene cluster, which is predicted to encode genes homologous to genes in the RK-682 biosynthetic gene cluster, such as a type I polyketide synthase, an Acyl-CoA

synthetase/dehydrogenase, an Acyl carrier protein, a FabH-like protein, and a FkbH-like protein. The aim of this work was to determine if the genes from the putative tetronate-like cluster in *S. scabies* are expressed under laboratory conditions. Currently, the metabolic product of this cluster has not been identified. To analyse the expression of the putative tetronate-like cluster in *S. scabies*, reverse transcriptase PCR (RT-PCR) was carried out on the total RNA isolated from the wild type *S. scabies*.

The *scab63021* gene has been proposed to be involved in the regulation of the tetronate-like biosynthetic gene cluster acting as a transcriptional activator. To prove this hypothesis, the attempt was made to disrupt the gene in *S. scabies* using PCR targeting protocol. Analysis of the wild type *S. scabies* and the putative *scab63021* mutant genomes has been carried out to allow for comparisons in the patterns of expression of the genes in the putative tetronate-like gene cluster.

The transcriptional analysis of the *S. scabies* wild type RNA revealed that the cluster is expressed. Transcripts of some genes (listed in section 5.2) were obtained and confirmed by sequencing. The PCR analysis and sequencing of the Δ*scab63021* mutant genomic DNA has confirmed the presence of the resistance cassette in the mutant genome. However, this did not enable us to conclude whether the cassette was in the correct position within the genome. The expression pattern does not illustrate any differences between the wild type strain and the Δ*scab63021* mutant indicating that *scab63021* does not have any effect on the expression of genes within the putative tetronate biosynthetic gene cluster.

Further work would be to carry out the comparative analysis of the *S. scabies* wild type strain and the Δ*scab63021* mutant culture supernatants via LC-MS to find the putative tetronate product. If there would be no differences between the metabolic profiles of the two strains, the Δ*scab63021* mutant could be constructed again by another method. Alternatively, the deletion of the *scab62941,* encoding for a putative PKS within the *S. scabies* tetronate-like gene cluster could be obtained and the metabolic profile of the wild type strain and the mutant could be compared again via LC-MS in search of the cryptic tetronate metabolite.

6.3 Significance

Among actinomycetes, members of the genus *Streptomyces* are considered to be the major producers of secondary metabolites (Challis and Hopwood, 2003). They have evolved the ability to biosynthesize multiple biologically active compounds. These compounds are not made during the vegetative growth phase and their production is usually switched on at a point when nutrients needed for a primary metabolism are depleted and during the aerial mycelium growth phase and during spore growth initiation. Their biological functions include, among the others, killing of competitive microorganisms, regulation of bacterial metabolism and metal ion uptake. It was argued that possible reasons for the biosynthesis of secondary metabolites may be the selectional advantages these compounds confer to the producing organisms such as the defence mechanism against the environmental stresses (Stone and Williams, 1992).

Siderophores are one class of secondary metabolites. They are produced by many bacteria to increase the solubility of iron from the soil. The analysis of several

Streptomyces species genomes were found to encode for the biosynthesis of multiple siderophores (Challis and Hopwood, 2003; Barona-Gomez et al., 2006), for example, *S. coelicolor* produces two hydroxamate siderophores: coelichelin and desferrioxamine. *S. avermitilis* produces desferrioxamine and is also predicted to produce the siderophore myxochelin (Ikeda et al., 2003). The reason for the production of more than one kind of iron-chelating metabolite is not exactly clear. However, it is known that many bacteria that do not produce desferrioxamine siderophore posses the uptake systems for it (Barona-Gomez et al., 2006). It was, therefore, hypothesized that streptomycetes, which produce desferrioxamine, also need to possess a biosynthetic gene cluster responsible for the production of another siderophore, which is not taken up into the cell by the ferroxamine uptake system (Challis and Hopwood, 2003).

Multiple pathways for the assembly of various siderophores were also found in several plant pathogenic microorganisms (Franza and Expert, 2012). It is known that iron is essential for the development of bacterial infections in mammalian hosts (Cox et al 1981). However, the role of iron uptake in the invasion of plants by pathogenic bacteria has only been investigated in recent years (Franza and Expert, 2012). These investigations have provided some evidence for the role of iron acquisition in the virulence of Gram-negative plant pathogenic microorganisms. For example, *Dickeya dadantii* a soft-rotting Gram-negative soil bacterium is a plant pathogen of many plant species, including some vegetables. The two siderophores that it produces- chrysobactin and achromobactin- are required for the development of infection symptoms in the plant tissues (Dellagi *et al*., 2005; Enard *et al*., 1988). Iron uptake is also required for infection by *Erwinia amylovora*, which causes fire blight in apple

flowers (Dellagi et al., 1998). *Pseudomonas syringae* pv. tabaci 6605 produces the siderophore pyoverdin, which is needed for development of wildfire disease on tobacco plants (Taguchi et al., 2010). As iron is often critical for the virulence of Gram-negative plant pathogens, it is suspected that siderophores are also important for iron uptake in pathogenic *Streptomyces* species. *Streptomyces scabies* causes disease in potatoes and root crops (Loria et al., 1998) and its genome encodes biosynthetic pathways directing biosynthesis of several siderophores, including pyochelin and desferrioxamines. The novel iron-chelating metabolite scabichelin was also isolated from this strain. Siderophore-mediated iron acquisition has not been yet confirmed to be a significant factor in pathogenicity of *S. scabies*. However, the investigations of the siderophore biosynthetic pathways in this strain will enable further elucidation of the role of iron acquisition in the virulence of plant pathogenic streptomycetes. Analysis of the *S. scabies* genome also unveiled the presence of a gene cluster encoding biosynthetic pathways for the production of secondary metabolites that are not associated with iron uptake but, which were reported to be present in other plant pathogens (Section 5). The investigation of these clusters and their metabolites will contribute to further elucidation of the molecular mechanism of *S. scabies* that are important for plant pathogenicity.

7. Materials and Methods

7.1 Materials

7.1.1 Enzymes, Chemicals and Equipment

Kits for DNA gel extraction and plasmid isolation from *E. coli* were purchased from Qiagen (USA). The kit for DNA ligation was obtained from Roche (Germany). Restriction enzymes were purchased from MBI Fermentas (Lithuania) and New England Biolabs (USA). Oligonucleotide primers were ordered from Sigma Genosys (USA). The gene synthesis was ordered from Genewiz. Taq polymerase was purchased from Fermentas and High Fidelity Polymerase from Roche. The ingredients for media preparation were purchased from Fisher BioReagents, Difco and Becton, Dickinson and co. The rest of chemicals were bought from Difco (USA), Sigma-Aldrich (USA), Fermentas (Lithuania), Fisher Scientific (UK), Biolab (New Zealand), Lonza (UK).

The concentration of DNA and RNA was measured on a NanoDrop ND-1000 Spectrophotometer. An Eppendorf Mastercycler Personal was used for PCR reactions. Agarose gel electrophoresis was performed on Bio-Rad Power Pac 300. A Bio-Rad Gene Pulser II with the Bio-Rad Pulse Controller Plus was used for electroporations. Thermo BioMate3 Spectrophotometer was used to check optical density of bacterial cell cultures. The aseptic work with *Streptomyces* was performed under Bassaire laminar flow hood. The microbial cell culture supernatants and extracts were analysed for secondary metabolite production by LC-MS (LC – Agilent 900, MS – Bruker HCT Plus) and Bruker MicroTOF mass spectrometer. For Southern blot hybridisation a UV Stratalinker 2400 from Stratagene was used for

DNA fixation to nylon membrane (Amersham). For hybridisation of probe to the DNA a TechHybridiser HB-1D was used.

7.1.2 General Solutions, Buffers and Antibiotics

All stock solutions (Table 7.1) were prepared according to Sambrook and Russel (2001). Distilled water was used for preparing all stock solutions. Autoclaving or filtration through a 0.2 µm filter (Stericon) was used to sterilise the solutions. Antibiotic stock solutions (Table 7.2) were dissolved in distilled water or other appropriate solution and sterilised by filtration through a 0.2 µm filter.

Table 7.1General stock solutions.

Chemical	**Stock solution**	**Solvent**
TBE (1000ML 5x)	Trizma base- 53g Boric acid- 27.5g EDTA- 10mM, pH 8	water
TE buffer	Tris –HCL pH 8 100mM EDTA – 10mM, pH 8	water
Tris-HCl (pH 8)	0.5M	Water, pH adjusted with HCL
Agarose gel loading dye	Glycerol 50% Bromophenol blue 0.1% Xylene cyanol 0.1%	water
RNAse solution	RNAse stock 50mg/ml	water

Table 7.2 Antibiotics stock solutions.

Antibiotic	**Stock sol. (mg/mL)**	**Solvent**	***E. coli* (µg/mL)**	***S. scabies* (µg/mL) overlay**	***S. scabies* (mg/plate) overlay**
ampicillin	100	water	100	-	-
apramycin	50	water	50	25	1.25
chloramphenicol	25	ethanol	25	-	-
hygromycin	50	water	50	25	1.25
kanamycin	50	water	50	100	5.0
nalidixic acid	25	0.3M NaOH	25	20	0.5

7.1.3 Bacterial Strains

Table 7.3 and table 7.4 list the bacterial strains used in this work.

Table 7.3 The *E. coli* strains used.

***Escherichia coli* strains**		
Name	**Genotype**	**Source**
DH5α	F- endA1 glnV44 thi-1 relA1 gyrA96 deoR nupG lacZdeltaM15 hsdR17	general cloning host (Woodcock et al., 1989), lab stock
MC1061	*araD139* Δ*(ara-leu)7696 galE15 galK16* Δ*(lac)X74 rpsL*(Strr) *hsdR2* (rK–mK+) *mcrA mcrB1*	general cloning host, lab stock
ET12567/pUZ8002	Non-methylating host for transfer of DNA into *Streptomyces* spp. plasmid pUZ8002*: tra*, KanR, *RP4 23;* Genome*: dam, dcm, hsdS,* CmR, TetR, JIC, Norwich	JIC, Norwich (Gust et al., 2004)
BW25113/pIJ790	Temperature sensitive strain of *E. coli*, should be grown on rich	Plant Bioscience

	medium containing chloramphenicol (25 μg/ml) at 30°C. Plasmid: pIJ790 *[oriR101], [repA101*(ts)*], araBp-gam-be-exo*. Chromosome: (Δ(*araD-araB*)*567*, Δ*lacZ4787*(::*rrnB-4*), *lacIp-4000*(lacIQ), λ-, *rpoS369*(Am), *rph-1*, Δ(*rhaD-rhaB*)568, *hsdR514;* Datsenko and Wanner, 2000).	Limited (Datsenko and Wanner, 2000)

Table 7.4 The *S. scabies* strains used

***Streptomyces scabies* strains**		**Source**
87.22	Wild type	Rose Loria Cornell University
Δ*scab1401*	87.22 *scab1401* null mutant	Rose Loria Cornell University
Δ*scab1471*	87.22 *scab1471* null mutant	Rose Loria Cornell University
Δ*scab1401-*Δ*scab1471*	87.22 *scab1401-scab1471* double mutant	Rose Loria Cornell University
Δ*scab desC*	*S. scabies* 87.22 *scab57931* null mutant	This study
Δ*scab desC comp*	*S. scabies* 87.22, *scab57931* null mutant complemented	This study
S. scabies W1000	*S. scabies scab8547*1::apr mutant	This study
Δ*scab63021*	*S. scabies* 87.22 *scab63021* ::apr mutant	Lab stock

7.1.4 Plasmids and Cosmids

Tables 7.5 list plasmids and cosmids used in this study.

Table 7.5 Plasmids and cosmids used.

Name	**Insert**	**Source/Reference**
pKC1132	Plasmid to create disruption in *scab85471*, derived from pOJ260 by deletion of the *Kpn*I-*Spe*I fragment, contains *oriT* RK2 for conjugation from *E. coli* to *Streptomyces*, *aac*(3)IV	Lab stock
pOSV556	Contains: *E. coli* origin of replication, *oriT*, hygromycin resistance gene, *ermE** promoter, pSAM2 integrase and corresponding attachement site.	Pernodet Group, University of Paris-Sud, France
pIJ773	Source of *oriT-apr* cassette for gene disruption (pBluescript KS(+), (*aac3(IV)*), *oriT*(RK2), FRT sites)	Gust et al. 2004
pUC57::*scab63021*	General cloning plasmid. Contains scab63021 and ampicillin resistance gene	Genewiz Inc.
S1757	Supercos1 derivative containing. scabies des cluster *scab57921* to *scab57981*genes	Rose Loria, Cornell University
S57	Supercos1 derivative containing *scab85251* to *scab85521*genes	Rose Loria, Cornell University

7.1.5 Primers

Oligonucleotide primers used in this study are listed in Table 7.6.

Table 7.6 Oligonucleotide primers used.

Name	**Sequence 5'→ 3'**	**Characteristic**
desCcomp_fw	aaagggaagcttaggaggcgcggtatgaccttca	Primers to complement *scab57931* gene
desCcomp_rv	gggaaactcgaggcgtgatgagccgctcg	
supcosseq_fw	aacaaatagggttccgc	Primers for sequencing ends of scabichelin gene cluster insert in Supercos 1
supcosseq_rv	gcctgtttttgtctctttcc	
replacement_fw	cagggcgcgcacccgctgccggtggatgaactgctgccgattccggggatccgtcgacc	PCR targeting primers to amplify *oriT-apramycin* cassette for *scab85471* disruption
replacement_rv	gttcgagtggaagaacaggccctcctggagcggggcgagtgtaggctggagctgcttc	
Tr_fw	cccggcccgcctccggcgctgccctccgcgccgcccggtattccggggatccgtcgacc	PCR targeting primers to amplify *oriT-apramycin* cassette for *scab63021* disruption
Tr_rv	cggcgtcgctctcgccggccgtggcggcggggcgggccctgtaggctggagctgcttc	
Test_fw	cgccatccacctcctcatcc	Used to confirm correct integration of S57::*oriT-apr* into the *S. scabies* 87.22 genome
Test_rv	aagcccggccagatccac	
desC_test_fw	gacgtcgcccgcaactac	Primers to test presence of *oriT-*

		apramycin cassette in the place of *scab57931*
desC_test_rv	gtgatgagccgctcgtgg	
Tr_test_fw	cgggcgtggaagtaaccg	Used to confirm correct integration of pUC57::*oriT-apr* into the *S. scabies* 87-22 genome
Tr_test_rv	ctgcacgagcgtgtcctg	
apr_test_fw	tacaccggaccttggagt	Used to confirm the presence of *oriT-apr* cassette in the *S. scabies* 87.22 genome
apr_test_rv	gaccgactggaccttccttc	

Table 7.7 RT-PCR primers used.

Name	**Sequence 5'→ 3'**	**Characteristic**
Scab62941_fw	gctgcgccggctggagga	primers to amplify internal 749 bp of *scab62941* gene
Scab62941_rv	tcttggtgcggttgctgctc	
Scab62951_fw	caacgaagagtggatcgacgt	primers to amplify internal 615 bp of *scab62951* gene
Scab62951_fw	tcgatggccttgtagtcctc	
Scab62971_fw	ctcggtcgtcacctatctg	primers to amplify internal 553 bp of *scab62971* gene
Scab62971_rv	gatgatccggtggtcgaag	
Scab62981_fw	gccgggaagctcaacctg	primers to amplify internal 617 bp of *scab62981* gene
Scab62981_rv	cagttggcggtgacctgc	
Scab62991_fw	ggcatcgtcagcgtcaag	primers to amplify internal 615 bp of *scab62991* gene
Scab62991_rv	cggagcagtgaggtgatcag	
Scab63021_int fw	gagtacccgttgcgggagagcc	primers to amplify internal 615 bp of *scab63021* gene
Scab63021_int rv	gatcaccgtgccctccgcgaag	
hrdB_fw	gcgtgtcgagcacggagtgcagc	primers to amplify internal 1317 bp of *hrdB* gene
hrdB_rv	ccgagtccgtctctgtcatggc	

7.1.6 Growth Media

Most growth media were prepared according to the manual by Kieser et al. (2000). All media were sterilised by autoclaving at 121°C.

7.1.6.1 Recipes for Liquid Media

LB Broth

Luria-Bertani powder	25g/L
H_2O	to 1L

2X YT

Difco Bacto Tryptone	16g
Difco Bacto Yeast extract	10g
NaCl	5g
H_2O	to 1L

SMM (Suplement Minimal Medium)

$MgSO_4$ .7H_2O (0.2M)	25 mL
TES Buffer (0.25M, pH 7.2)	100 mL
NaH_2PO_4+ K_2HPO_4(50 mM each)	10 mL
Trace Element Solution	1 mL
Casaminoacids (20%)	10 mL
H_2O	804 mL
Glucose (20%)	50 mL

Each solution was sterilised by autoclaving or filtering and then added to sterile water.

Trace Element Solution

$ZnSO_4 .7 H_2O$ 0.1 g/L

$FeSO_4 .7 H_2O$ 0.1 g/L

$MnCl_2 .4 H_2O$ 0.1 g/L

$CaCl_2 .6 H_2O$ 0.1 g/L

NaCl 0.1 g/L

H_2O to 1 L

This solution is sterilised by filtration.

Iron Deficient Medium (IDM; Barona-Gomez et. al., 2004)

Salt concentration

K_2SO_4 2 g

K_2HPO_4 3 g

NaCl 1 g

NH_4Cl 5 g

H_2O to 1 L

Trace element concentration for IDM

Thiamine 2 mg/L

$ZnSO_4.7H_2O$ 2 mg/L

$CuSO_4$ (0.5 mg/mL) 10 µL

$MnSO_4.H_2O$ (3.5mg/mL) 10 µL

$MgSO_4 .H_2O$ 80 mg/L

Other solutions

$CaCl_2$	10 g/l
Glucose	250 g/l
Yeast extract	0.5%

These solutions were filter sterilised after preparation.

The 1 litre of the IDM solution was stirred with 50 g of Chelex-100 Sodium form (Sigma Aldrich) in a plastic flask, covered with parafilm and left on the magnetic stirrer over night to let the Chelex resin remove the traces of iron. The Chelex was then filtered off on a funnel with filter paper (Whatmann) and the trace element solutions were added in the following amounts: 100 µl of thiamine and 100 µl of $ZnSO_4.7H_2O$ stock solution, 10 µl of $CuSO_4$ stock solution, 10 µl of $MnSO_4.H_2O$ and 80mg of $MgSO_4.7H_2O$. The medium was then sterilised by autoclaving. Before starting the culture $CaCl_2$, glucose and yeast extract stock solutions were added (500 µl of stock solution per 50 ml media).

7.1.6.2 Recipes for Solid Media

Solid agar recipes used for *Streptomyces* growth.

OBB (Oat Bran Broth; Johnson et al., 2007)

Oatmeal	18g
Bacto Agar	18g
H_2O	to 1L

Trace element added after autoclaving.

SFM (Mannitol soya flour medium)

Soya flour	20g
Mannitol	20g
Difco Bacto Agar	20g
H_2O	to 1L

R5 medium

Ingredients used before autoclaving

Sucrose	103 g
K_2SO_4	0.25 g
$MgCl_2.6H_2O$	10.12 g
Glucose	10 g
Difco Casaminoacids	1 g
Difco Yeast Extract	5 g
TES Buffer	5.73 g
Agar	22 g
H_2O	to 1L

All ingredients from the list above were mixed and medium was autoclaved. At the time of use the medium was melted and the ingredients from the list below were added.

After autoclaving ingredients/1L

KH_2PO_4 (0.5%)	10 mL
$CaCl_2.2H_2O$ (5M)	4 mL
L-proline (20%)	15 mL
NaOH (1N)	7 mL
Trace Element Solution	2 mL

Solid agar recipes used for *E. coli* growth.

LB Agar

Luria-Bertani powder	25 g/L
Difco Bacto Agar	15 g/L
H_2O	to 1L

DNA (Difco nutrient agar)

Difco Nutrient Agar	4.6 g
H_2O	to 1L

7.2 Growth, Manipulation and Storage of *E. coli*

The procedures used for growth, manipulation and storage of *E. coli* were taken from Sambrook and Russel (2001).

7.2.1 Growth Conditions of *E. coli*

E. coli was cultured over night in LB broth (180 rpm, 37°C) or on LB agar plates at 37°C. Antibiotic selection was used for plasmid-containing cells at the concentrations listed in Table 7.2.

7.2.2 Preparation of Electrocompetent *E. coli* Cells

E. coli was grown over night at 37°C in 10 ml LB shaking at 180 rpm. 100 µl of *E. coli* overnight culture was then inoculated into 10 ml LB and grown for 3-4 hours shaking at 200 rpm to an OD_{600} of ~ 0.4. The cells were then recovered by centrifugation at 3000 rpm for 10 min at 4°C. After decanting the medium the cell pellet was resuspended by mixing in 10 ml of ice-cold 10% glycerol. The cells were then centrifuged again and resuspended in 5 ml of ice-cold 10% glycerol, centrifuged, decanted and resuspended in the remaining ~100 µl 10% glycerol.

7.2.3 Transformation of Electrocompetent *E. coli* Cells

About 100 ng of DNA were added to 100 µl of *E. coli* cell suspension in a single transformation. Electroporation was done in a 0.2 cm ice-cold electroporation cuvette using a BioRad GenePulser II set to: 200 Ω, 25 µF and 2,5 kV and the expected time constant of 4.5 – 4.9 ms. Immediately after electroporation, 1 ml of ice cold LB was added to shocked cells. They were then incubated shaking for 1h at 37°C and spread onto LB agar plates containing appropriate antibiotic for the selection of transformants.

7.2.4 Storage of E. coli Cells

E. coli overnight LB cultures were mixed with 70% glycerol to a final concentration of 20% and stored at -78 °C for a long term.

7.3 Growth, Manipulation and Storage of *Streptomyces*

The procedures used for growth, manipulation and storage of *Streptomyces* were described by Kieser et al. (2000). The protocol for genomic DNA extraction was taken from Pospiech and Neumann (1995).

7.3.1 Spore Stock Collection from Agar Plates

S. scabies spores were grown on oat bran broth (OBB) agar. Spores were collected from each plate by adding 3 ml of sterile water and scrapping the surface of the mycelia with a sterile loop. The spore suspension was then filtered through sterile cotton wool and centrifuged for 10 min at 3000 rpm. The supernatant was poured off and the spore pellet was resuspended in approximately 500 µl of the remaining water. The stock was then made by adding similar volume of 50% sterile glycerol to the spore suspension. The spores were stored at -20ºC.

7.3.2 Liquid Cultures for Genomic DNA Extraction

S. scabies spores for genomic DNA extraction were grown in 50 ml of TSB medium at 30ºC, 180 rpm for 72 hours.

7.3.3 DNA Transfer from *E. coli* to *S. scabies* by Conjugation

10µl of *E. coli* ET12567/pUZ8002 glycerol stock was added to 10 ml of a liquid LB medium containing kanamycin (25 µg/ml) and chloramphenicol (25 µg/ml) and the

culture was grown at 37ºC over night.Then, 100 µl of over night culture was added to 10 ml LB and incubated for 6 hours with the same antibiotics and at the same conditions as above. After 6 hours the electrocompetent cells were prepared according to the procedure used earlier (section7.2.2), transformed with the *oriT*-resistance cassette containing cosmid or plasmid clone (section 7.2.3) and selected for the incoming cosmid or plasmid by plating on LB agar with kanamycin (25 µg/ml) and chloramphenicol (25 µg/ml) as well as antibiotic appropriate to the resistance of the cassette. One of the transformant colonies was inoculated into 10 ml LB liquid medium containing the same antibiotics as those used on LB plates and grown overnight at 37ºC. Then, 500 µl of overnight culture was inoculated into 50 ml of fresh LB plus antibiotics as above and grown for 6 h at 37°C, shaking at 180 rpm to an OD_{600} of 0.4. After 6 hours the cells were washed twice with 50 ml of LB to remove antibiotics (might inhibit *Streptomyces*), and resuspended in ~2 ml of LB. While washing the *E. coli* cells, 500µl (approximately 10^9) of freshly collected *Streptomyces* scabies 87.22 spores were added to 0.5ml 2 × YT broth, heat shocked at 50°C for 20 min, then cooled for 1 min on ice. Next, ~2 ml of *E. coli* cell suspension were mixed with 500 µl heat-shocked spores in 500 2xYT and then centrifuged briefly. The supernatant was then partially removed and the pellet was resuspended in the c.a. 500 µl residual liquid. This liquid was then plated out without diluting on several SFM agar + 10mM MgCl2 plates (without antibiotics) and incubated at 30°C for 16-20 h. The next day each plate was overlayed with 1 ml of water containing 0.5 mg nalidixic acid (20 µl of 25 mg/ml stock) to selectively kill *E. coli*. The antibiotics in the appropriate concentrations were also added for the

selection of the exconjugant *Streptomyces* containing the plasmid or cosmid of interest.

7.4 Extraction and Manipulation of DNA

7.4.1 Isolation of Genomic DNA from *S. scabies*

The *S. scabies* DNA was isolated according to the method by Pospiech and Neumann (1995).

The cells of *S. scabies* 87.22 strain were grown in 50ml of TSB medium at 30°C for three days. The cells were harvested by centrifugation (10 min, 3000 rpm) and resuspended in 5 ml of SET (75 mM NaCl, 25 mM EDTA, 20 mM Tris, pH 7.5). Lysozyme was added to the cells to a concentration of 1 mg/ml and incubated at 37°C for 0.5 h. Next, l/10 volume of 10% SDS and 0.5 mg/ml of proteinase K were added and the cells were incubated at 55°C with an occasional inversion for 2 h. After that l/3 volumes of 5 M sodium chloride and one volume of chloroform were added and the cells were incubated at room temperature for 0.5 h with frequent inversion. The cells were then centrifuged at 4500 *g* for 15 min; the aqueous phase was transferred to a new tube using a blunt-ended pipette tip. The DNA was precipitated by adding 1 volume of isopropanol and gently inverting the tube. The precipitated DNA was then transferred into a microfuge tube and rinsed with 70% ethanol, dried on air and dissolved in 1ml of TE buffer. The amount of DNA obtained was between 1-3 mg.

7.4.2 Plasmid/Cosmid Isolation from *E. coli*

The *E.coli* cosmid or plasmid was isolated according to the method by Sambrook and Russell (2001).

5 ml of liquid LB medium with added antibiotics was inoculated with *E. coli* cells and they were grown overnight at 37°C, 180 rpm. The cells were then centrifuged and the cell pellet was resuspended in 100 µl of solution I (50 mM Tris/HCl, pH 8; 10 mM EDTA). Then, 200 µl solution II (200 mM NaOH; 1% SDS) was added to it and cells were mixed by inverting the tube several times. Then, 150 µl of solution III (3 M potassium acetate, pH 5.5) were added to the tube and it was mixed by inverting several times. The tube was centrifuged at a maximum speed in a microcentrifuge for 5 min at room temperature. Then, the supernatant was transferred to a new tube and 400 µl of phenol-chloroform-isoamyl alcohol (25:24:1) was added, vortexed 2 min and centrifuged at a full speed in a microcentrifuge for 5 min, at room temperature. The upper phase was transferred to a new tube and 600 µl of ice cold isopropanol was added to it. The tube was left on ice for 10 min to precipitate the DNA, and the sample was centrifuged for 10 min at a full speed in a minicentrifuge at 4°C. The supernatant was removed and the pellet was washed with 200 µl 70% ethanol, spinned in a microcentrifuge for 5 min at room temperature and left for 30 min at room temperature to air-dry. The pellet was then resuspended in a 30 µl of sterile water with RNAse.

7.4.3 Digestion of DNA with Restriction Enzymes

Digestions by restriction enzymes were prepared following the manufacturer instructions. Restriction enzymes always occupied 10% of the total volume and

appropriate buffer also occupied 10% (all buffers were 10 x concentrated). If BSA had to be added, it also occupied 10% of the total volume. The rest of the total volume consisted of a DNA solution to be digested and, if necessary, water. Digestions of plasmid/cosmid were carried out in a total volume of 10 µl.

7.4.4 Ligation of DNA into Plasmids

DNA to be ligated into plasmid vectors was purified using the Fermentas gel extraction kit and eluted with 10 µl of distilled water. The concentration of vector and insert was then estimated by running 1µl of each DNA sample on an agarose gel along with a molecular size marker of a known DNA concentration. A molar ratio of vector: insert of 1:4 was generally used for ligation reactions. The ligation reactions were carried out in a 20 µl of a total volume for 30 min using Rapid Ligation kit from Roche.

7.4.5 Agarose Gel Electrophoresis

1x TBE buffer was used to prepare and run agarose gels. Different percentage agarose gels were prepared (0.8%, 1%, 1.5%) depending on the size of DNA fragments. 2 µl of Gel Red dye (Biotium) was added to the gel to visualize DNA under UV light. Electrophoresis was carried out for 40 min at 110 V. A. 1kb DNA molecular size marker (Fermentas) was used as a size standard.

7.5 PCR Methods

The High Fidelity Polymerase (Roche) was used to amplify a gene for PCR targeting or for cloning. The Taq polymerase (Fermentas) was used for analysis of constructs

or mutants. The PCR reactions were carried out in 50 µl total volume. The PCR reaction conditions are displayed in table 7.8.

Table 7.8 PCR reaction conditions.

Solution	Volume	Concentration
Primers (100 pmoles/µl)	1 µl each	100 pmoles each
Template DNA	1 µl	~ 50 ng
Buffer (10x)	5 µl	1x final concentration
dNTPs (5 mM)	2 µl each	100 µM each
DMSO (100%)	2.5 µl	5%
DNA polymerase (2.5 U/µl)	0.5 µl	1.25 Units
Distilled water	37µl	-
Total volume	50 µl	-

7.5.1 General PCR Method

PCR conditions to amplify inserts for cloning and for analysis of constructs and mutants are displayed below.

1) Denaturation: 94°C, 2 min
2) Denaturation: 94°C, 45 s
3) Primer annealing: 55-67°C, 45 s } 35 cycles
4) Extension: 72°C, 90-240 s
5) Final extension: 72°C, 10 min

A primer annealing temperature is chosen to get the maximum copies of a product and the highest specificity. It is usually between 55-67°C. The extension time is chosen according to the rule 1 min of extension per 1kb of product.

7.5.2 PCR Amplification of the Gene Replacement Cassette.

A PCR reaction (Gust et al., 2002) with designed primers was carried out to amplify the cassette from the digested pIJ773 vector. The conditions of the reaction are as follows:

1) Denaturation:	94°C, 2 min	
2) Denaturation:	94°C, 45 s	10 cycles
3) Primer annealing:	50°C, 45 sec	
4) Extension:	72°C, 90 sec	
5) Denaturation:	94°C, 45 sec	25 cycles
6) Primer annealing:	55°C, 45 sec	
7) Extension:	72°C, 90 sec	
8) Final extension:	72°C, 5 min	

7.6 PCR Targeting in *S. scabies* 87.22

PCR targeting technique was used to create in-frame deletions in *S. scabies* genome (Gust et al., 2002).

7.6.1 Design of PCR Primers

Two targeting primers were designed for the deletion in each gene according to the procedure described in section 3.1. The targeting primers are displayed in Table 7.6 (Section 7.1.5).

7.6.2 Purification of the Resistance Cassette

Approximately 3 μg of plasmid DNA pIJ773 was digested with 50 U *Eco*RI (Roche) and 50 U *Hin*dIII (Roche), in 1 x buffer (Roche), in a 50 μl reaction for 1.5 hour at 37° C. A 2938 bp vector fragment and a 1400 bp for the cassette were generated. The

digest was then run on a gel and then the band for the resistance cassette was cut out and purified by Qiagen gel extraction kit. The purified fragment was stored in a 30 µl of a sterile water, at –20°C.

7.6.3 PCR Amplification of the Extended Resistance Cassette

The purified resistance cassette was amplified by PCR with long targeting primers (Table 7.6) using the reaction mixture shown in Table 7.9. The Expand High Fidelity polymerase (Roche) was used in PCR reactions under the conditions described in section 7.5.2. 5 µl of the PCR reaction was analysed by electrophoresis on a 1% agarose gel for the presence of the cassette with the size of ~1476 bp. The rest of the PCR product was purified from the reaction mixture using the Qiagen PCR purification kit following the manufacturer's instructions and eluted from the column with 30 µl of water.

7.6.4 Introduction of Cosmid Clones into *E. coli* BW25113/pIJ790 by Electroporation

pIJ790 vector contains the resistance marker *cat* (chloramphenicol resistance) and a temperature sensitive origin of replication (requires 30°C for replication). *E. coli* BW25113/pIJ790 cells were grown overnight at 30°C in a 10 ml LB containing chloramphenicol (25 µg/ml). 100 µl of *E. coli* BW25113/pIJ790 from the overnight culture were inoculated in a 10 ml LB containing 20 mM $MgSO_4$ (200 µl of 1M stock) and chloramphenicol (25 µg/ml) and grown for 4 hours at 30°C, shaking at 180 rpm to an OD_{600} of ~ 0.4. The competent cells were then prepared according to the section 7.2.2 and the 100 µl of the cell suspension was mixed with ~ 100 ng (1-2

µl) of a cosmid DNA. Electroporation was carried out as described in section 7.2.3 and the cells were incubated on an orbital shaker at 180 rpm for 1h at 30°C. The culture was spread onto two plates of LB agar containing ampicillin (100 µg/ml), kanamycin (50 µg/ml) and chloramphenicol (25 µg/ml) and incubated overnight at 30°C. The next day, the transformant colonies containing cosmid were observed. One isolated colony was transferred into 5 ml of liquid LB containing antibiotics as above and incubated overnight at 30°C. This culture was then used as a pre-culture for generating competent cells to be transformed with the extended resistance cassette.

7.6.5 PCR Targeting of Cosmids

10 ml of LB containing ampicillin (100 µg/ml), kanamycin (50 µg/ml) and chloramphenicol (25 µg/ml) was inoculated with 1% of the overnight culture of *E. coli* BW25113/pIJ790 and the cosmid. 100 µl of 1M L-arabinose stock solution was added to the culture to induce the *red* genes. Then, the culture was grown for 4 hours at 30°C shaking, at 180 rpm to an OD600 of ~ 0.4. Competent cells were prepared as described in section 7.2.2 and the 50 µl of the cell suspension was mixed with ~ 100 ng (2 µl) of the extended resistance cassette PCR product. Electroporation was carried out (section 7.2.3) and the cells were incubated shaking at 180rpm for 1h at 37°C (the temperature above 30°C causes degeneration and loss of pIJ790 plasmid). The culture was spread onto two plates of LB agar containing ampicillin (100 µg/ml), kanamycin (50 µg/ml) and apramycin (50 µg/ml) and incubated overnight at 37°C. After 16 hours growth at 37°C different colony sizes were observed. The small colonies were false positives and the larger the sizes of a colony, the more copies of

the mutagenised cosmids were present. Several large colonies were inoculated in 5 ml LB liquid cultures containing ampicillin (100 µg/ml), kanamycin (50 µg/ml) and apramycin (50 µg/ml) and grown at 37°C, shaking at 180 rpm to a cell density OD_{600} ~ 0.1 – 0.3, for 6 h. After 6 hours, the cosmid DNA was isolated and tested by restriction analysis and PCR.

About 2 µg (10 µl) of the wild type cosmid clone DNA and 2 µg (10 µl) of the mutagenised cosmid (with introduced extended apramycin cassette and OriT) were digested, each with *Bam*HI or *Sac*I restriction enzymes in a total volume of 30 µl. Both digestions were left for a 3 hours at 37°C and then 5 µl of each digestion was analysed on an agarose gel (1% agarose). The digested wild type and mutagenised cosmid gave distinct patterns of bands. The expected digestion pattern for the mutagenised cosmid was confirmed by an *in silico* prediction using program, such as Clone Manager or NEB cutter. If the digestion pattern agreed with the program prediction then the integration of the disruption cassette into cosmid was also confirmed using PCR with suitable test primers (Table 7.6) priming within ~ 100 bp upstream and downstream of the region, which recombined with the PCR product of the extended *OriT*-apramycin resistance cassette. The PCR reaction was performed under the conditions listed in the section 7.5.2. The mutagenised cosmid was introduced into a non-methylating *E. coli* host ET12567/pUZ8002 by electroporation (section 7.3.3).

7.6.6 Transfer of the Mutant Cosmids into *Streptomyces scabies*

The procedure for conjugal transfer of the engineered cosmid from *E. coli* ET12567/pUZ8002 into *S. scabies* is described in section 7.3.3. After several days of

incubation the grey, hydrophobic *Streptomyces* colonies were present on each plate. Several colonies were picked and transferred onto DNA plates (Difco Nutrient Agar) containing nalidixic acid (25µg/ml) and apramycin (50µg/ml) with and without kanamycin (50µg/ml). Double cross-over exconjugants were kanamycin sensitive (the cosmid backbone containing kanamycin resistance gene has been lost during cross over) and apramycin resistant (cassette contains apramycin). The kanamycin sensitive clones were picked from DNA plates and plated out on oat bran broth (the solid medium used for growing of *S. scabies*). The spores were then collected; genomic DNA isolated and analysed for the desired mutation by PCR (section 7.5.1) with the test primers (section 7.1.5) and also by Southern blot hybridisation (section 7.7).

7.7 Southern Blot Hybridisation

DIG High Prime DNA Labelling and Detection Starter Kit (Roche) was used to carry out Southern blot hybridisation.

7.7.1 Probe Labeling

1µg of cosmid DNA (5 µl) was added to a final volume of 16 µl of autoclaved miliQ distilled water and denatured by boiling at 100°C for 10 min. The sample was then rapidly cooled on ice, mixed with 4 µl of DIG High Prime solution and incubated over night in 37°C water bath. Then, the labeling reaction was stopped by adding 2 µl of 0.2 M EDTA (pH 8.0) and then by heating to 65°C for 10 min.

7.7.2 Genomic DNA Digestion

20 µg of genomic DNA for each strain was digested with *Bam*HI or *Sac*I over night at 37°C. The samples were then loaded on 0.7% agarose gel and run over night at 25

V. The gel was then exposed to UV light for 5 minutes to help denature the DNA. Then, the gel was rinsed in distilled water and washed twice 15 minutes in denaturation buffer (2 x 15 min shaking at room temperature). After denaturation the gel was rinsed again in distilled water and twice with neutralisation buffer (2x 15 shaking).

7.7.3 Capillary Transfer and DNA Fixation

A glass plate was placed across a plastic tray containing 20xSSC buffer. One sheet of 3mm Whatmann filter paper soaked in 20xSSC buffer was placed on the glass plate. The ends of the filter paper touched the buffer in the container. On the top of this filter paper another soaked piece of paper was put and on that the treated gel with the wells side down. The piece of a nylon membrane with a positive charge cut to the size of the gel was placed on the top of it. The gel was wrapped around on the edges with a tape to prevent 'short cirquiting.' Then, three pieces of Whatmann 3 mm paper the size of gel were placed on the top of a nylon membrane; after that, the 5 cm high layer of paper towels and finally on that a glass plate with 2 litres flask half-filled with water as a weight. The movement of the DNA to the membrane was carried out by capillary transfer over night. The next day the membrane was removed from the gel, placed on the piece of Whatmann 3mm paper soaked in 10xSSC buffer and UVcross-linked to fix the DNA (auto-crosslinking on UV Stratalinker 2400 from Stratagene).

7.7.4 Hybridisation

The membrane with the DNA transferred to it was prehybridised in a hybridisation tube in a 20 ml of DIG Easy Hyb solution (DIG Easy Hyb granules mixed with 64 ml of sterile water for 10 min at 37°C) rotating for 2 hours at 42°C. The membrane was then incubated in the hybridisation solution over night rotating at 42°C. The hybridisation solution was prepared by adding the 20 µl of denatured DIG labeled DNA probe (5µg/µl) into 7 ml of preheated DIG Easy Hyb solution. The probe was denatured by boiling for 5 min at 98°C and then rapidly cooled on ice. The next day the hybridization solution was removed and the membrane was washed 2x 5 minutes in 50 ml preheated 2x SSC, 0.1% SDS rotating at 25°C; then 2x 15 min in 50 ml preheated 0.5x SSC, 0.1% SDS at 58°C. The membrane was then washed in 30 ml of washing buffer for 5 min at room temperature. Then, it was incubated at 50 ml of blocking solution at room temperature; under constant rotation in the hybridisation oven. After 30 minutes the blocking solution was replaced with a 30 ml of the fresh blocking solution, mixed with a 3 µl of the anti-DIG antibody conjugate and the membrane was incubated with it for another 30 min. Then, the antibody solution was removed and the membrane was washed twice in 50 ml of washing buffer (2x15 min rotating) and then equilibrated for 5 min in 20 ml of a detection buffer.

7.7.5 Detection

The wet membrane was placed with the DNA side up in a hybridisation bag and 2 ml of chemiluminescence substrate CSPD (from DIG High Prime kit, Roche) was applied onto it. The membrane was then immediately covered with the second sheet of the development bag so that the CSPD substrate was spread evenly without an air

bubbles over the membrane. The edges of the hybridisation bag were then sealed and the membrane was incubated in the development folder in the dark. After 5 min the membrane was placed inside a new hybridisation bag, sealed and incubated at 37°C for 15 min to enhance the luminescence reaction. The membrane was then exposed to film over night. The film was then developed according to standard procedures.

7.8 RNA Methods

7.8.1 Total RNA Isolation from *S. scabies*

The wild type *S. scabies* were grown on an oat bran broth plates at 30°C for 3, 5, 7 and 14 days. After the incubation, the mycelia from each plate were collected by a plastic loop and ground in a mortar with the addition of a liquid nitrogen. The frozen ground mycelium was transferred into a 1.5 ml Eppendorf tube containing 350 µl of resuspension buffer RA1 from NucleoSpin RNA II kit (Macherey-Nagel) and 3.5 µl of mercapthoethanol. The sample was then vortexed vigorously for 3 minutes to help the cell wall breakage. The sample was processed further using NucleoSpin RNA II kit. The RNA was finally eluted on the column with 60 µl of RNAse free distilled sterile water. The RNA was always handled in gloves. All pipette tips and Eppendorf tubes were autoclaved twice. The bench area and pipettes were wiped with the RNAse inhibitor agent (RNaseZAP, Sigma Aldrich). The RNA was analysed using agarose gel elecrophoresis on a 1% agarose gel. 1µl of RNA was mixed with the 4 µl of the loading dye (Fermentas) and run on a gel with 50 V current for 1hour. The RNA was then visualised under the UV light. The presence of the two sharp bands and the third lower and more faded was indicating that the RNA is not degraded and

it is suitable for downstream applications, such as PCR. The RNA samples were then treated with DNAse (TurboDnase, Ambion) to remove any traces of genomic DNA. They were then analysed using standard PCR method (section 7.5.1) with Taq polymerase to ensure that no genomic DNA was present in the samples. The purified RNA was then used as a template in reverse transcriptase PCR reactions (section 7.8.2) to analyse expression of the genes from the putative tetronate biosynthetic gene cluster.

7.8.2 Reverse Transcriptase PCR Method

The Qiagen One Step PCR Kit was used for performing of PCR reactions. Polymerase from that kit is a mix of reverse transcriptase (RT) polymerases (Omniscript Reverse Transcriptase, Sensiscript ReverseTranscriptase) and the Taq DNA polymerase (Hotstar Taq polymerase). The RT polymerase converts messenger RNA (mRNA) into a complementary DNA (cDNA) in the initial part of the PCR reaction carried out at 50°C in 30 minutes. In the next step, the temperature of the reaction was raised to 95°C to activate the Taq DNA polymerase, deactivate the reverse transcriptases and denature the cDNA. The cDNA template was then amplified by the Taq polymerase in a standard PCR reaction.

The PCR reaction conditions are displayed in the following table:

Table 7.9 RT-PCR reaction conditions

Solution	**Volume**	**Amount**
Primers (1x)	0.5-1 µl each	0.5-1 µM each
Template RNA	1-2 µl	1pg-2µg
5x QIAGEN OneStep RT-PCR Buffer	10 µl	1x final concentration
dNTPs (10 mM)	2 µl each	400 µM each
QIAGEN OneStep RT-PCR Enzyme Mix (5 U/µl)	2 µl	2.5 Units
RNase-free water	variable	-
Total volume	50 µl	-

Thermal cycler conditions for RT-PCR:

1) Reverse transcription: 50°C, 30min

2) Initial PCR activation step: 95°C, 15 min

3-step cycling:

1) Denaturation: 94°C, 1 min

2) Annealing : 50-68°C, 45 s

3) Extension: 72°C, 1 min/kb

35 cycles

4) Final extension 72°C, 10 min

7.9 Growth of *Streptomyces scabies*; Extraction of Metabolites and Analysis of Metabolite Production

7.9.1 Pyochelin

7.9.1.1 Culturing of *S. scabies* in Iron Deficient Medium

S. scabies 87.22 and and the Δ*Scab1401* and Δ*Scab1401*–Δ*Scab1471* mutants were cultured in Iron Deficient Medium for a detection of pyochelin. 50 ml of sterilised IDM in each 250 ml sterile plastic Erlenmeyer flask were inoculated with 10 µl of fresh spore stock of each *S. scabies* pyochelin mutant and wild type *S. scabies* strain. The liquid cultures were incubated at 30°C shaking at 180 rpm for 4-5 days. 1 ml of each cell culture was then taken to 1.5 ml Eppendorf tubes and centrifuged in a table microfuge at 13000 rpm for 5 minutes at room temperature. Supernatant was placed in 1 ml HPLC tubes. The samples were analysed using LC-MS.

7.9.1.2 LC-MS Analysis

Liquid chromatography –Mass Spectrometry (LC-MS) was used to analyse the supernatants of the *S. scabies* 87.22 and the Δ*Scab1401* and Δ*Scab1401*–Δ*Scab1471* mutants. LC-MS analysis of the reaction mixture was carried out using Eclipse XDB-C18 column (150 x 4.6mm, 5µm, Agilent) connected to an Agilent 1100 HPLC instrument. The outflow was connected via a splitter (10% flow to MS, 90 % flow to waste) to a Bruker HCT+ mass spectrometer fitted with an electrospray source in positive ion mode with parameters: nebulizer flow 40 psi, dry gas flow 10.0 L/min, dry temperature 300 ºC, capillary – 4 kV, skimmer 40V, capillary exit 106 V, ion charge control target (ICC) 100,000, spectral averages 3. Absorbance was

monitored at 210 nm and 280 nm. The elution profile is displayed in the table below, with a 1 ml/min flow rate. The retention time for pyochelin was 23 minutes.

Table 7.10 Gradient elution profile used in LC-MS analyses of pyochelin production.

Time (min)	**Water + 0.1% formic acid**	**Methanol + 0.1% formic acid**
0	100	0
5	100	0
25	0	100
30	0	100
35	100	0

7.9.1.3 Extraction of Pyochelin and High Resolution Mass Spectrometry

S. scabies 87.22 and the Δ*Scab1401* and Δ*Scab1401*–Δ*Scab1471* mutants were grown in an iron-deficient medium (Barona-Gomez et al., 2004) for 4 days. The cultures were centrifuged for 10 min at 4000 r.p.m. at 4°C. The supernatants were lyophilised,

resuspended in 5 ml of methanol and analysed using a Dionex Ultimate 3000RS instrument fitted with Supelco Acentis Express C18 column (15062.1 mm, 2.7 mm); coupled to a Bruker MaXis mass spectrometer with parameters: ESI in positive mode; full scan 50–1500 m/z; End plate offset, 2500 V; Capillary, 24500 V; Nebulizer gas (N2), 1.6 bar; Dry gas (N2), 8 l min21; Dry Temperature, 180 uC. The elution profile is displayed in the table. The flow rate was 0.2 ml per minute. Absorbance was monitored at 210 and 280 nm. The High Resolution Mass spectrometry analyses were performed by our collaborator Dr. Lijiang Song

Table 7.11 Gradient elution profile used in High Resolution MS analyses of pyochelin production.

Time (min)	**Water + 0.1% formic acid**	**Methanol + 0.1% formic acid**
0	100	0
5	100	0
30	0	100
35	0	100
40	100	0
55	100	0

Pyochelin and enantio-pyochelin (ent-pyochelin) standard solutions (both 10 µl, 10 µg/ml in methanol) and concentrated culture supernatant from the Δ*Scab1401* mutant were analysed using a homochiral stationary phase using an Agilent 1100 HPLC coupled to a Bruker HCT plus mass spectrometer with parameters: ESI in positive mode; full scan 50–1500 m/z; End plate offset, 2500 V; Capillary, 24500 V; Nebulizer gas (N2), 40 p.s.i.; Dry gas (N2), 10 l min21; Dry Temperature, 300 uC. The eluents and elution conditions are described in Table 7.11. The flow rate was 1 ml/min. All High Resolution Mass Spectrometry analyses for pyochelin and scabichelin, as well as, chiral HPLC analyses, were carried out with the assistance of our collaborator on the project Dr. Lijiang Song.

7.9.2 Scabichelin

7.9.2.1 LC-MS Analysis

S. scabies 87.22 and W1000 were grown in iron deficient medium (Barona-Gomez et al., 2004) according to the same protocol as in section 7.9.1.1. 1 ml of supernatants from each of 50 ml cultures of *S. scabies* 87.22 and W1000 were analysed by high resolution LC-MS on a Sigma Ascentis Express column (C18, 150 x 2.1mm, 2.7μm) attached to a Dionex 3000RS UHPLC coupled to a Bruker MaXis ESI-Q-TOF mass spectrometer. The mobile phases were water containing 0.1% formic acid (eluent A) and methanol containing 0.1% formic acid (eluent B). A gradient of 5% B to 100% B in 15 minutes was used with a flow rate of 0.2ml/min. Absorbance at 210nm was monitored. The mass spectrometer was operated in a positive ion mode with a scan range 50-2,000 *m/z*. Source conditions were: end plate offset at -500V; capillary at -4500V; nebulizer gas (N_2) at 1.6bar; dry gas (N_2) at 8L/min; dry Temperature at 180 °C. Ion tranfer conditions were: ion funnel RF at 200Vpp; multiple RF at 200Vpp; quadruple low mass at 55 m/z; collision energy at 5.0 eV; collision RF at 600 Vpp; ion cooler RF at 50-350 Vpp; transfer time at 121μs; pre-Pulse storage time at 1 μs.

7.9.3 Desferrioxamines

7.9.3.1 Growing *S. scabies* in Supplemented Minimal Medium (SMM) and LC-MS Analysis

50 ml of sterilised SMM in each 250 ml sterile plastic Erlenmeyer flask was inoculated with 10 µl of fresh spore stock of *S. scabies* Δ*desC* mutant and wild type *S. scabies* strain. The liquid cultures were then incubated at 30°C incubator, shaking at 180 rpm for 3-4 days. 1 ml of each cell culture was then taken to 1.5 ml Eppendorf tubes and 2µl of $FeCl_3$ was added to them. The samples were incubated for 30 min at room temperature and then centrifuged in a table microfuge at 13000 rpm for 5 minutes at room temperature. Supernatant was taken to 1 ml HPLC tubes. The samples were then analysed using LC-MS. The LC-MS equipment and analysis conditions were same as described for pyochelin LC-MS analysis (Section 7.9.1.2). The absorbance was monitored at 435 nm.

References

Angert, E. R.; *Nature Rev Microbiol* **2005**, *3*, 214-224.

Ankenbauer, R.; Sriyosachati, S. and Cox, C. D. *Infect Immun* **1985**, *49*, 132–140.

Austin, M.B. and. Noel, J. P. *Nat Prod Rep* **2003**, *20*, 79-110.

Barona-Gomez F.; Lautru S.; Francou F. X.; Pernodet J. L.; Leblond P.; Challis G. L. *Microbiol* **2006**, *152*, 3355-3366.

Barona-Gomez, F.; Wong, U.; Giannakopulos, A. E.; Derrick, P. J. and Challis, G. L. *J Am Chem Soc*, **2004**, *126*, 16282–16283.

Barry S. M.; Kers J. A.; Johnson E. G.; Song L.; Aston P. R.; Patel B.; Krasnoff S. B.; Crane B. R.; Gibson D. M.; Loria R.; Challis G. L. *Nat Chem Biol* **2012**, *8*, 814-816.

Bell, K. S., Sebaihia, M., Pritchard, L., Holden, M. T., Hyman, L. J., Holeva, M. C., Thomson, N. R., Bentley, S. D., Churcher, L. J., Mungall, K., Atkin, R., Bason, N., Brooks, K., Chillingworth, T., Clark, K., Doggett, J., Fraser, A., Hance, Z., Hauser, H., Jagels, K., Moule, S., Norbertczak, H., Ormond, D., Price, C., Quail, M. A., Sanders, M., Walker, D., Whitehead, S., Salmond, G. P., Birch, P. R., Parkhill, J., and Toth, I. K. *Proc. Natl. Acad. Sci. U.S.A.*, **2004**, *101*, 11105-11110.

Berti A. D. and Thomas M. G. *J. Bacteriol* **2009,** *191***,** 4594-4604.

Bickel H.; Bosshardt R.; Gaumann E.; Reusser P.,Vischer E., Voser W., Wettstein A., Zahner A. *Helv Chim Acta* **1960**, *43*, 2118–2128.

Bignell, D. R. D.; Huguet-Tapia J. C.; Joshi M. V.; Pettis, G. S.; Loria, R. *Antonie van Leeuwenhoek* **2010,** *98*, 179–194. (1)

Bignell D. R. D., Seipke R. F., Huguet-Tapia J. C.,Chambers A. H., Parry R. J., Loria R *The Am. Phytopathol Soc*, **2010**, *23*, 161 – 175. (2)

Brickman, T. J. and McIntosh, M. A. *J Biol Chem* **1992,** *267*, 12350-12355.

Bruner S. D.; Weber T.; Kohli R. M., Schwarzer D., Marahiel M. A., Walsh C. T. and M. Stubbs T., Structure **2002**, *10*, 301-310.

Bukhalid, R. A.; Gibson D. M.; Crane, B R.; Loria R. *Nature* **2004**, *429*, 79–82.

Cairo, G., Bernuzz,i F. and Recalcati, S. *Genes and Nutriton* **2006**, *1*, 25-40.

Castignetti, D. *Curr Microbiol* **1997**, *34*, 250–257.

Carbonetti, N. H. and Williams, P. H. *Infect. Immun* **1984**, *46*, 7–12.

Cartinhour, S. W. and Shuler, M. L. *Appl Environ Microbiol* **2009**, *75*, 2720–2726.

Challis G. L., *ChemBioChem* **2005**, *6*, 601-611.

Challis, G. L. *J Am Chem Soc* **2004**, *126*, 16282–16283.

Challis, G. L. and Hopwood, D. A. *Proc Natl Acad Sci U S A* **2003,** *100*, 14555-61.

Challis, G. L. and Lautru, S. *Microbiol* **2004**, *150*, 1629-1636.

Challis, G. L. and Ravel, J. *FEMS Microbiol Lett* **2000,** *187*, 111-114.

Chan, Y. A., Podevels, A. M., Kevany, B. M. and Thomas, M. G. *Nat Prod Rep* **2009**, *26*, 90-114.

Cox, C. D. *Infect Immun* **1982**, *36*, 17-23.

Cox, C. D; Rinehart, K. L.; Moore, M. L.; Cook, J. C. *Proc Natl Acad Sci U S A* **1981**, *78*, 4256-4260.

Crone, W. J. K.; Leeper, F. J.; Truman, A. W. *Chem Sci* **2012**, 11, 3121-3127.

Cruz, A.T.; Cazacu, A. C.; Allen, C. H. *J Clin Microbiol* **2007**, *45*, 1989-1992.

Deeth, R.J. *Coordin Chem Rev* **2001**, *212*, 11-34.

Dellagi, A.; Brisset, M., N.; Paulin, J., P. and Expert, D. *Mol Plant Microbe Interact* **1998**, *11*, 734–742.

Dellagi, A.; Rigault, M.; Segond, D.; Roux, C.; Kraepiel, Y.; Cellier, F.; Briat, J.F.; Gaymard, F. and Expert, D. *Plant J.* **2005**, *43*, 262–272.

Demydchuk, Y.; Sun, Y.; Hong, H.; Staunton, J. B.; Leadley, P. F. *Chem biochem* **2008,** *9*, 1136-1145.

Donadio, S.; Monciardini, P.; Sosio, M. *Nat Prod Rep* **2007,** *24*, 1073-109.

Donadio, S.; Sosio, M.; Lancini, G.; *Appl Microbiol Biotechnol* **2002**, *60*, 377–380.

Drose, S.; Bindseil, K. U.; Bowman, E. J.; Siebers, A.; Zeeck, A.

and Altendorf, K. *Biochem, N Y* **1993**, *32*, 3902–3906.

Drose, S.; Boddien, C.; Gassel, M.; Ingenhorst, G.; Zeeck, A.; Altendorf, K. *Biochemistry* **2001**, *40*, 2816–2825.

Dunbar, K.L.; Melby, J. O.; Mitchell, D. A. *Nat Chem Biol* **2012**, *8*, 569.

Enard, C., Diolez, A. and Expert, D. *J Bacteriol*, **1988**, 170, 2419–2426.

Fernandez-Moreno, M. A.; Caballero, J. L.; Hopwood, D. A.; Malpartida, F. *Cell* **1991**, *66*, 769-780.

Finking, R. and Marahiel, M. A. *Ann Rev Microbiol* **2004**, *58*, 453-488.

Franza, T. and Expert D. *Mol Plant Pathol* **2012**, 1-10.

Fry, B. A.; Loria, R. *Physiol Mol Plant Path* **2002**, *60*, 1-8.

Garg, R. P. and Parry, R. J. *Microbiol* **2010**, *156*, 472-483.

Gehring, A. M.; Bradley, K. A. and Walsh, C. T. Biochem **1997**, *36*, 8495–8503

Gehring, A. M.; Mori I.; and Walsh, C. T. *Biochem* **1998**, *37*, 2648–2659.

Gibson, D. M. *Mol Microbiol* **2009**, *73*, 409-18.

Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L. *Chem Sci 2012*

Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc Natl Acad Sci U S A* **2003,** *100*, 1541-6.

Gust, B.; Kieser, T.; Chater, K. F. REDIRECT(c) Technology: PCR-targeting system in *Streptomyces coelicolor*. 2002.

Gust, B., O'Rourke, S., Bird, N., Kieser, T. and Chater, K. F. (2003b). Recombineering in Streptomyces coelicolor. Norwich, UK: The John Innes Foundation.

Hamaguchi, T.; Sudo, T.; Osada, H. *FEBS Lett* **1995**, *18*, 54-8.

Haydock, S.F.; Appleyard, A. N.; Mironenko, T.; Lester, J.; Scott, N.; Leadlay, P. F. *Microbiol* **2005**, 151, 3161–3169.

Healy, F. G.; Krasnoff, S. B.; Wach, M.; Gibson, D. M.; Loria, R. *J Bacteriol* **2002**, *184*, 2019–2029.

Heinrichs, D. E. and Poole, K. *J Bacteriol* **1993**, *175*, 5882–5889.

Heinrichs, D. E. and Poole, K. *J Bacteriol* **1996**, *178*, 2586–2592.

Hoegy, F.; Lee, X.; Noe, S.l; Rognan, D.; Mislin, G. L. A.; Reimmann, C. and Schalk, I. J. *J Biol Chem* **2009**, *284*, 14949-14957.

Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. and Omura, S. *Nat. Biotechnol.* **2003**, *21***,** 526–531.

Imbert, M.; Bechet, M. and Blondeau, R. *Curr Microbiol* **1995**, *31*, 129–133.

Ino, A. and Murabayashi, A. Tetrahedron **2001**, *57*, 1897–1902

Johnson, E. G.; Krasnoff, S. B.; Bignell, D. R.; Chung, W. C.; Tao, T.; Parry, R. J.; Loria, R.; Johnson, K. B. and Stockwell, V. O. *Annu Rev Phytopathol* **1998**, *36***,** 227-248.

Johnson, E. G.; Joshi, M. V.; Gibson, D. M. and Loria, R. *Physiol Mol Plant Pathol* **2007**, *71*, 18-25.

Kadi, N.; Oves-Costales, D.; Barona-Gomez, F.; Challis, G.L. *Nat Chem Biol* **2007**, *3*, 652-656.

Katz, E., and Demain, A.L. *Bacteriological Rev* **1977**, *41*, 449-474.

Kämpfer, P. The Prokaryotes: a Handbook on the Biology of Bacteria, Springer: Berlin, 2006.

Kellenberger, J. L.; Ph.D Thesis. ETH, 1997.

Kers, J.A.; Wach, M.J.; Krasnoff, S.B.; Widom, J.; Cameron, K.D.; Bukhalid, R.A.; Gibson, D.M.; Crane, B.R.; Loria, R. *Nature* **2004** 429:79–82

Kers, J. A.; Wach, M. J.; Krasnoff, S.B.; Widom, J.; Cameron, K. D.; Kershaw, J. K.; Lennard, N. J.; Jones, M.; Jones, S. J. M. and Solenberg, P. J. *Chem Biol* **1998**, *5*, 155-162.

Kessler, N.; Schuhmann, H.; Morneweg, S.; Linne, U. and Marahiel, M. A. *J Biol Chem* **2004**, *279*, 7413-7419.

Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. Practical *Strepto-*

myces Genetics. 2 ed.; The John Innes Centre Fundation: Norwich, 2000.

Kim, B. J.; Park, J. H.; Park, T. H.; Bronstein, P. A.; Schneider, D. J., Cartinhour, S. W.; Shuler, M. L. Appl Environ Microbiol **2009**, *75*, 2720-6.

Kinashi, H; Someno, K.; Sakaguchi, K. *J Antibiot, (Tokyo)* **1984**, *37*, 1333–1343

King, R. R.; Calhoun, L. A. *Phytochemistry* **2009**,*70*, 833-41.

King, R.R.; Lawrence, C. H.;Calhoun, L. A.; Ristaino, J. B. *J Agric Food Chem* **1994**, *42*, 1791–1794.

King, R.R.; Lawrence, C. H.; Clark, M.C.; Calhoun, L.A. *J Chem Soc Chem Commun* **1989**, *13*, 849–850.

Kobayashi, Y. M.;. Icchioka, T.; Hirose, K.; Nagai, A.; Matsumoto, H.; Matsui, H.; Hanaki, R.; Masuma, Y.; Takahasi Y. and Omura S. *Bioorg Med Chem Lett* **2010**, *20*, 6116

Kraemer, S. M. *Aqua Sci* **2005**, 66, 3–18.

Labeda D.P., Goodfellow M., Brown R., Ward A. C., Lanoot B., Vanncanneyt M., Swings J., Kim S.B., Liu Z., Chun J., Tamura T., Oguchi A., Kikuchi T., Kikuchi H., Nishii T., Tsuji K., Yamaguchi Y., Tase A., Takahashi M., Sakane T., Suzuki K.I., Hatano K. *Antonie Leeuwenhoek* **2012**, *101*, 73-104.

Laitala-Leinonen, T.; Lowik, C.; Papapoulos, S. and Vaananen, H. K. *J Cell Sci* **1999**, *112*, 3657–3666.

Lambert, D. H. and Loria, R. *Int J Syst Bacteriol*, **1989**, *39*, 387-392.

Lautru, S. and Challis, G.L. *Microbiol* **2004**, *150*, 1629-1636.

Lautru, S.; Deeth, R. J.; Bailey, L. M. and Challis, G. L. *Nat Chem Biol* **2005**, *1*, 265-269.

Lawrence, C. H.; Clark, M. C.; King, R. R. *Phytopathology*, **1990**, *80*, 606–608.

Li, N.; Zhang C.; Li, B.; Liu, X.; Huang, Y.; Xu, S.; Gu, L. *J Biol Chem* **2012**, *287*, 8912-8919.

Lichon, V.; Khachemoune, A. *Am J Clin Dermatol* **2006**, *7*, 315-21.

Liu, J.; Quinn, N.; Berchtold, G. A. and Walsh, C. T. *Biochem* **1990**, *29*, 1417–1425.

Loria, R.; Bignell, D. R.; Moll,S.; Huguet-Tapia, J. C.; Joshi, M. V.; Johnson, E. G.; Seipke, R. F.; Gibson, D. M. *Antonie van Leeuvenhoek*, **2008**, *94*, 3-10.

Loria, R. and Kempter, B. A. *Plant Dis* **1986,** *70*, 1146–1148.

Loria, R.; Kers, J.; Joshi, M. *Annu Rev of Phytopathol* **2006**, *44*, 16.1-16.19.

Luo, M.; Fadeev, E. A.; Groves, J. T. *Nat Chem Biol* **2005**, *1*, 149-153.

Marfey, P.*Carisberg Res Commun* **1984**, *49*, 591-596.

Marsh, H. V.; Evan, H. J.; Matrone, G. *Plant Physiol* **1963**, *38*, 638–642.

McMurry, J.; Begley, T. The Organic Chemistry of Biological Pathways. Roberts and Company Publisher: Englewood, CO, 2005.

Mellmann, I.; Fuchs, R. and Helenius, A. *A Rev Biochem* **1986**, *55*,663–700.

Meyer, J. M.; Neely, A.; Stintzi, A.; Georges, C.; Holder, I. A. *Infect Immun* **1996**, *64*, 518–523.

Michel, L.; Gonzalez, N.; Jagdeep, S.; Nguyen-Ngoc, T. and Reimmann, C. *Mol Microbiol* **2005**, *58*, 495–509.

Miethke, M., and Marahiel, M.A. *Microbiol Mol Biol Rev* **2007**, *71*, 413-451.

Miller M. J. *Chem Rev*, **1989**, ***7***, 1563-1579.

Muller, G. and Raymond, K. N. *J Bacteriol* **1984**, *160*, 304–312.

Narva, K. E. and Feitelson, J. S. *J Bacteriol* **1990**, *172*, 326-333.

Natsume, M.; Komiya, M.; Koyanagi, F.; Tashiro, N.; Kawaide, H.; Ab,e H.; *J Gen Plant Pathol* **2005,** 71, 364–369.

Natsume, M.; Ryu, R.; Abe, H. *Ann Phytopathol Soc Jpn* **1996**, *62*, 411–413.

Natsume, M.; Yamada, A.; Tashiro, N.; Abe, H. *Ann Phytopathol Soc Jpn* **1998**, *64*, 202–204.

Natsume, M.; Komiya, M.; Koyanagi, F.; Tashiro, N.; Kawaide H.;Abe, H. *J Gen Plant Pathol* **2005**, *71*, 364–369.

Neilands, J. B. J Biol Chem **1995**, 270, 26723–26726.

Nelson, N. (1995). Molecular and cellular biology of F- and VATPase.

In *Organellar Poton-ATPases* (ed. N. Nelson), pp. 1–27. New York: Springer.

Omura, S.; Imamura, N.; Hinotozawa, K.; Otaguro, K.; Hashimoto, K.; Nakagawa, A. *J Antibiot* **1982**, *35*, 1632–1637.

Otaka, T. and Kaji, A. *J. Biol Chem* **1976**, *251*, 2299.

Otaka, T. and Kaji, A. *FEBS Lett* **1981**, *123*, 173.

Otaka, T. and Kaji, A. *FEBS Lett* **1983**, *153*, 53.

Oves-Costales D.; Kadi N.; Challis G. L.; Chem Comm **2009**, *43*, 6530-6541.

Patel, P.; Song, L. and Challis, G. L. *Biochemistry* **2010**, *49*, 8033-8042.

Patel, H. M. and Walsh, C. T. *Biochem* **2001**, *40*, 9023–9031.

Pohlmann, V. and Marahiel, M.A. *Org Biomol Chem* **2008**, *6*, 1843-1848.

Pospiech, A., and Neumann, B. *Trends Genet* **1995**, *11*, 217-218.

Rangaswamy, V.; Jiralerspong, S.; Parry, R.; and Bender, C. L. *Proc Natl Acad Sci U S A* **1998**, *95*, 15469-15474.

Raymond, K. N.; Dertz, E. A.; Kim, S. S. *Proc Natl Acad Sci U S A* **2003**, *100*, 3584-3588.

Reichert, J.; Sakitani, M.; and Walsh, C. T. *Prot Sci* **1992**, *1*, 549–556.

Reimann, C.; Patel, H. M.; Serino, L.; Barone, M.; Walsh, C. T.; Haas, D. *J Bacteriol* **2001**, *183*, 813-820.

Reimmann, C.; Patel, H. M.; Walsh, C. T. and Haas, D. *J Bacteriol* **2004**, *186*, 6367–6373.

Rausch, C.; Weber, T.; Kohlbache,r O.; Wohlleben, W. and Huson, D. H. *Nucleic Acids Res* **2005**, *33*, 5799-5808.

Rusnak, F.; Faraci, W. S. and Walsh, C. T. *Biochemistry* **1989**, *28*, 6827–6833.

Rusnak, F.; Liu, J.; Quinn, N.; Berchtold, G. A. and Walsh, C. T. *Biochemistry* **1990**, *29*, 1425–1435.

Quadri, L. E. N.; Keating, T. A.; Patel, H. M. and Walsh, C. T. *Biochemistry* **1999**, *38*,14941–14954.

Sakaitani, M.; Rusnak, F.; Quinn, B.; Frigo, T. B.; Tu, C.; Berchtold G. A. and Walsh, C. T. *Biochemistry* **1990**, *29*, 6789–6798.

Sambrook, J.; Russell, D. W. Molecular Cloning - A Laboratory Manual. 3 ed.; Cold Spring Harbor Laboratory Press: New York, 2001.

Schlegel, K.; Taraz, K. and Budzikiewicz, H. *BioMetals* **2004**, *17*, 409–414.

Schupp, T.; Waldmeier, U. and Divers, M. *FEMS Microbiol Lett* **1987**, *42*, 135–139.

Schupp, T.; Toupet, C. and Divers, M. *Gene* **1988**, *64*, 179–188.

Seipke, R.F.; Song, L.; Bicz, J.; Laskaris, P.; Yaxley, A.M.; Challis, G.L.; Loria, R. *Microbiol* **2011**, *157*, 2681-93.

Seki-Asano, M.; Okazaki, T.; Yamagishi, M.; Sakai, N.; Hanada, K.; Mizoue, K. *J Antibiot* **1994**, *47*, 1226–1233.

Serino, L.; Reimmann, C.; Baur, H.; Beyeler, H.; Visca, P. and Haas, D. *Mol Gen Genet* **1995**, *249*, 217-228.

Serino, L.; Reimmann, C.; Visca, P.; Beyeler, M.; Chiesa, V. D. and Haas, D. *J Bacteriol* **1997**, *179*, 248–257.

Serino, L.; Reimann, C.; Visca, P.; Beyeler, M.; Chiesa, D. V. and Haas D. *J Bacteriol* **2003**, *179*, 248-257.

Seyedsayamdost, M. R.; Traxler, M. F.; Zheng, S.-L.; Kolter, R. and Clardy, J. *JACS* **2011**, *133*, 11434-11437.

Shaw-Reid, C. A.; Kelleher, N. L.; Losey, H. C.; Gehring, A. M.; Berg, C.; Walsh, C. T. *Chem Biol* **1999**, *6*, 355-400.

Shimamura, H.; Gouda, H.; Nagai, K.; Hirose, T.,; Ichioka, M.; Furuya, Y.; Kobayashi, Y.; Hirono, S.; Sunazuka, T. and Ōmura, S. *Angew Chem Int Ed* **2009**, *48*, 914.

Shoji, J.; Sakazaki, R.; Hattori, T.; Matsumoto, K.; Uotani, N.; Yoshida, T. *J Antibiot* **1989**, *42*, 1729-33.

Smith, D. J.; Earl, A. J. and Turner, G. *EMBO J* **1990**, *9*, 2743-2750.

Sriyosachati, S. and Cox, C. D. *Infect Immun* **1986**, *52*, 885–891.

Stone, M. J. and Williams, D. H. *Mol Microbiol U S A* **1992,** *6*, 29-34.

Sun, Y.; Hahn, F.; Demydchuk, Y.; Chettle, J.; Tosin, M.; Osada, H.; Leadlay, P. F. *Nat Chem Biol* **2010**, *6***,** 99-101.

Sun, Y.; He, X.; Liang, J.; Zhou, X.; Deng, Z. *Appl Microbiol Biotechnol* **2009**, *82*, 303.

Strohl, W. R. *Nucleic Acids Res* **1992,** *20*, 961-974.

Taguchi, F.; Suzuki, T.; Inagaki, Y.; Toyoda, K.; Shiraishi, T. and Ichinose, Y. *J Bacteriol* **2010**, *192*, 117–126.

Takase, H.; Nitanai, H.; Hoshino, K. and Otani, T. *Infect Immun* **2000**, *68*, 1834-1839.

Tanaka, A.; Takano, Y.; Ohnishi, Y.; Horinouchi, S. *J Mol Biol* **2007**, 369, 322-333.

Thariath, A.; Socha, D.; Valvano, M. A. and Viswanatha, T. *J Bacteriol* **1993**, *175*, 589–596.

Thomas, M. S. BioMetals **2007**, *20*, 431–452.

Wageningen, A. A.; Kirkpatrick, P. N.; Williams, D. H., Harris B. R; Kershaw, J.K.; Lennard, N. J.; Jones, M.; Jones, S. J. M. and Solenberg, P. J. *Chem Biol* **1998**, *5*, 155-162.

Waisvisz, J. M.; van der Hoeven, M. G.; van Peppen, J.; Zwennis, W. C. M. *J Am Chem Soc* **1957**,*79*, 4520.

Walsh, C. T.; Gehring, A. M.; Weinreb, P. H.; Quadri, L. E.; Flugel, R. S. *Curr Opin Chem Biol* **1997,** *1*, 309-15.

Warner, P. J.; Williams, P. H.; Bindereif, A. and Neilands, J. B. *Infect Immun* **1981**, 33, 540–545.

Wietzorrek, A.; Bibb, M.; *Mol Microbiol* **1997**, *25*, 1181–1184.Wodzinski, R. S.; Umholtz, T. E.; Beer, S. V. *J Appl Bacteriol* **1994**, *76***,** 22-29.

Yaxley, A. M.; Ph.D Thesis. University of Warwick, 2009.

Youard, Z. A.; Mislin, G. L. A.; Majcherczyk, P. A.; Schalk, I. J. and Reimmann, C. *J Biol Chem* **2007**, *282*, 35546-35553.